DOCUMENT RESUME

ED 380 777 CS 012 076

TITLE	NAEP Reading Revisit: An Evaluation of the 1992 Achievement Levels Descriptions.
INSTITUTION	American Coll. Testing Program, Iowa City, Iowa.
SPONS AGENCY	National Assessment Governing Board, Washington, DC.
PUB DATE	Feb 95
CONTRACT	ZA9003001
NOTE	79p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; Grade 4; Grade 8; Grade 12; *Item Analysis; *Reading Achievement; Reading Research; *Reading Tests; *Test Interpretation; *Test Items; Test Use
IDENTIFIERS	*Educational Issues; National Assessment of Educational Progress

ABSTRACT

A study was designed to provide recommendations regarding the use of the achievement levels set in 1992 for reporting National Assessment of Educational Progress (NAEP) reading results in 1994 and in future NAEP reading assessments. Two procedures were used: the Item Difficulty Categorization (IDC) procedure involved an evaluation of the achievement levels descriptions (ALDs) via a statistical categorization of items; and the Judgmental Item Categorization (JIC) evaluated the ALDs via a judgmental item mapping. A total of 56 panelists were trained in the reading framework and achievement levels set in 1992. The high correspondence between the judgmental classification and the statistical classification provided compelling evidence that the achievement levels descriptions communicate clearly and accurately with respect to student performance. Further, the recommendations developed by the panelists involved in the two evaluation methods were quite similar. The fact that the recommendations made by the two sets of panelists were similar and confirming of the achievement levels seemed a sufficiently positive outcome to support the use of the achievement levels for reporting the 1994 results. To the extent that panelists recommended changes in the achievement levels descriptions, it was to increase the requirements for the Basic level descriptions, based on actual student performance. Contains nine tables of data. Appendixes present a list of observers, panelists, and staff; nomination material; achievement levels descriptions and procedures for evaluating; the agenda; examples of IDC and JIC lists; grade-level organizational and format suggestions for ALDs; and an alternative method of computing "hits." (RS)

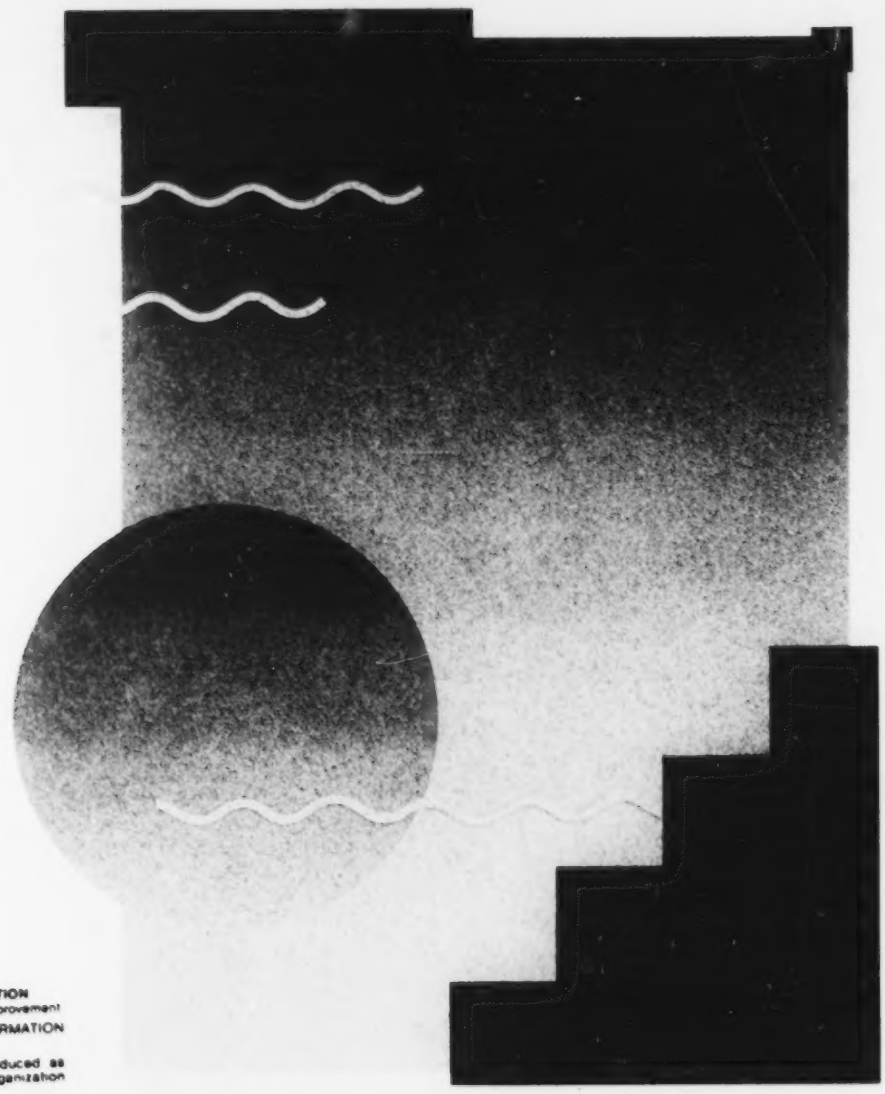
****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ACT

ED 380 777

NAEP Reading Revisit
An Evaluation of the 1992 Achievement Levels Descriptions

Presented by American College Testing
February 1995



CS012076

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

NAEP Reading Revisit
An Evaluation of the 1992 Achievement Levels Descriptions

Presented by American College Testing
February 1995

The work for this report was conducted by The American College Testing Program under modification of contract ZA9003001 with the National Assessment Governing Board.

Copyright © 1995 by The American College Testing Program. All rights reserved.

TABLE OF CONTENTS

Overview 1

Selection of Panelists 2

Training of Panelists 6

Procedures for Evaluating Achievement Levels 7
- Item Difficulty Categorization Procedure 7
- Judgmental Item Categorization Procedure 9

Results 10
- Results for the IDC Procedure 10
- Results for the JIC Procedure 11
- Joint Evaluation Results (Recommendations and Observations) 17
- Correspondence Between the IDC and JIC Procedures 21

Conclusion 27

References 28

NAEP Reading Revisit
An Evaluation of the 1992 Achievement Levels Descriptions

Overview

On October 15, 1994, three panels were convened in St. Louis, Missouri to evaluate the achievement levels descriptions for the 1992 National Assessment of Educational Progress (NAEP) in Reading. The National Assessment Governing Board (NAGB) had invited American College Testing (ACT) to submit a proposal for procedures to help address concerns raised by the National Academy for Education (NAE) in their evaluation of the NAEP Trial State Assessment, including the achievement levels-setting (ALS) process (NAE: 1994). ACT submitted the proposal to NAGB's achievement levels committee during the March 1994 meeting, and it was approved by NAGB during the May 1994 meeting.

The NAE recommended that "NCES and NAGB not report the 1992 NAEP results by achievement levels." (p. 132) They also recommended "against the use of the 1992 achievement levels in . . . reading as baselines against which to make comparisons in future assessments..." (p. 134)

Several suggestions for the re-examination of the 1992 reading achievement levels were put forth. These suggestions generally involved methods that generated new achievement levels descriptions without setting new achievement levels. ACT's proposal was explicitly *not* to develop new achievement levels descriptions. Rather, this study was designed to provide recommendations regarding the use of the achievement levels set in 1992 for reporting NAEP reading results in 1994 and in future NAEP reading assessments.

Panel members were, therefore, given the task of recommending whether the 1992 NAEP Reading achievement levels descriptions in 1992 should be used for reporting results for the 1994, and subsequent, Reading Assessment. Representatives from agencies involved with various aspects of the NAEP were also present to observe the evaluation process. (Please see Appendix A for the lists of staff, panelists, and observers.)

Two procedures were implemented for this study to evaluate the 1992 NAEP Reading Achievement Levels. Each grade level panel was divided into two groups: one group for each procedure. The first procedure was implemented to address the question of whether specific statements in the achievement levels descriptions (ALDs) are supported by performance on items in the NAEP for students scoring within the ranges of the achievement levels. The second procedure was implemented to examine the extent to which the skills and knowledge described in the achievement levels descriptions (ALDs) correspond to those measured by the assessment. The conjunction of these two procedures provides a review of the achievement levels descriptions in a systematic and thorough way that addresses the issue of whether the achievement levels can be used to report results on the 1994 Reading NAEP.

Selection of Panelists

Panelists for the Reading Revisit were nominated by coordinators and liaisons of the state affiliates of the National Council of Teachers of English (NCTE) and by state reading specialists and directors/coordinators of reading curriculum. A total of 165 individual nominators were identified from these two sources to serve as nominators. Each nominator was invited to submit up to 10 nominations: 6 teachers and 4 other (nonteacher) educators.

Guidelines to nominators specified that panelists should have (1) direct familiarity with knowledge and skills of students at a specific grade level, (2) training and experience in teaching reading, (3) involvement in professional organizations, and (4) other relevant qualifications. Each panel was to consist of 60% K-12 classroom teachers and 40% other educators. In addition, efforts were made to select panelists representing different regions of the country, and different gender and ethnic groups. (Please see Appendix B for the nomination materials.)

Of the 146 nominees, 83 were K-12 teachers and 63 were other educators. A total of 69 nominees were contacted; of these, 56 agreed to participate. Demographic characteristics (e.g., teaching status, gender, race/ethnicity) are provided in Tables 1, 2, and 3.

Table 1 provides distributions for the pool of nominees. Table 2 shows the distributions for panelists who were selected and invited to participate. Some of those panelists were unable to participate, so additional nominees were invited to be panelists. The data in Tables 2 and 3 show very similar distributions with respect to the sex and race/ethnicity of panelists invited to participate. The representation of male panelists was disturbingly low, and the selection process nearly depleted the nominee pool.

The panelists who served were very well qualified for the task. The credentials were most impressive in terms of professional activities and accomplishments.

Table 1

Percentages of Nominees, by Grade Level, Teaching Status, Gender, and Race/Ethnicity

	Teaching Status		
	TR	NT	Total
Grade 4	21	22	43
Grade 8	32	20	52
Grade 12	30	21	51
Total	83	63	146

	Gender		
	Male	Female	Total
Grade 4	6	37	43
Grade 8	3	49	52
Grade 12	8	43	51
Total	17 (12%)	129 (88%)	146

	Race/Ethnicity		
	White	Minority	Total
Grade 4	33	10	43
Grade 8	42	10	52
Grade 12	40	11	51
Total	115 (79%)	31 (21%)	146

Table 2

Percentages of Selected Nominees, by Grade Level, Teaching Status, Gender, and Race/Ethnicity

	Teaching Status		
	TR	NT	Total
Grade 4	14	7	21
Grade 8	13	12	25
Grade 12	15	8	23
Total	42 (61%)	27 (39%)	69

	Gender		
	Male	Female	Total
Grade 4	5	16	21
Grade 8	2	23	25
Grade 12	7	16	23
Total	14 (20%)	55 (80%)	69

	Race/Ethnicity		
	White	Minority	Total
Grade 4	15	6	21
Grade 8	20	5	25
Grade 12	19	4	23
Total	54 (78%)	15 (22%)	69

Table 3

Percentages of Panel, by Grade Level, Teaching Status, Gender, and Race/Ethnicity

	Teaching Status		Total
	TR	NT	
Grade 4	11	7	18
Grade 8	11	7	18
Grade 12	12	8	20
Total	34 (61%)	22 (39%)	56

	Gender		Total
	Male	Female	
Grade 4	5	13	18
Grade 8	1	17	18
Grade 12	5	15	20
Total	11 (20%)	45 (80%)	56

	Race/Ethnicity		Total
	White	Minority	
Grade 4	13	5	18
Grade 8	15	3	18
Grade 12	17	3	20
Total	45 (80%)	11 (20%)	56

Training of Panelists

Prior to meeting, panelists were asked to review a set of materials including: (1) the *Reading Framework*, (2) *Descriptions of Reading Achievement Levels for Basic, Proficient, and Advanced Performance on the 1992 NAEP*, (3) a description of the Procedures for Evaluating Achievement Levels Set for the 1992 Reading NAEP, and (4) a summary version of the design document for setting 1994 NAEP achievement levels in Geography and U.S. History[1]. These materials were provided to help prepare panelists for the tasks of the study. (Please see Appendix C for some of these materials.)

Strict security arrangements were followed during the process to ensure that all materials were accounted for at all times. All participants signed and submitted a security agreement prior to reviewing materials.

The Reading Revisit involved a series of activities that occurred during October 15 and 16, 1994. (Please see Appendix D for the Agenda.) ACT and NAGB staff provided panelists with information regarding the background and purposes of NAEP, the development of the NAEP Reading Framework, the roles of the principal agencies and organizations involved in NAEP, the achievement levels-setting process implemented in 1992, and the subsequent efforts by both NAGB and NCES regarding the 1992 Reading NAEP. Panelists were given an overview of the process; their roles and tasks were described for them; and they were told how they would be trained and assisted to perform their roles and tasks.

To prepare panelists for the task of evaluating the reading achievement levels, each panelist was administered a form of the reading assessment for their grade level. Testing was conducted under timed conditions similar to those experienced by students. After completing the test, panelists were given scoring keys and rubrics for scoring their own examinations. This was intended to facilitate their understanding of the items, the "correct" answers, and the scoring methodology. Panelists also became familiar with the general content covered by the assessment, the time constraints imposed, and the general level of difficulty of the assessment.

Panelists were trained in the reading framework and achievement levels set in 1992 to help them form a common understanding. Charles Peters, a member of the reading framework consensus panel, described the reading framework to panelists. He provided a detailed analysis of the achievement levels descriptions, related them to the reading framework, and provided examples of assessment items.

---

[1] This document was sent to the panelists so that they would have some understanding of how the achievement levels were set. A letter explained that the document was to inform panelists about the general process, even though the Achievement Levels-Setting procedures for the 1994 NAEP were somewhat different from those that were implemented in 1992.

Following the presentation of the ALDs, a training session was conducted for the purpose of helping panelists reach a common understanding of the ALDs. Two exercises were implemented for this purpose. In the first exercise, each panelist used his or her understanding of the ALDs to estimate student performance on different types of items. One block of items was used for this exercise from the 1994 assessment for each grade level. This provided the opportunity to train panelists using items consistent with the framework that would not later be used in the main "revisit" task. The panelists were instructed to determine the skills and knowledge called for in each item and to try to match them to those in the achievement levels descriptions. Each panelist shared his or her conclusions with other panelists. Reasons for agreements and disagreements were discussed to help arrive at a common understanding of the meaning of the achievement levels descriptions.

In the second exercise, panelists were asked to determine the level of performance exhibited in sample responses to open-ended questions. Panelists were shown a sample of student responses to each extended response item in the 1992 assessment for their grade level. Panelists were asked to determine the level of performance (Basic, Proficient, Advanced) exhibited in each response. The concept of borderline performance was explained to panelists to facilitate their decisions.

All panelists participated in these exercises as members of one of the three grade level panels.

Procedures for Evaluating Achievement Levels

Two procedures were used to evaluate the 1992 achievement levels. The Item Difficulty Categorization (IDC) procedure involved an evaluation of the achievement levels descriptions *via* a statistical categorization of items. The other procedure, called the Judgmental Item Categorization (JIC) procedure, was used to evaluate the ALDs *via* a judgmental item mapping. Panelists judged which achievement level matched the knowledge and skill required by each item. Each grade level panel was divided into two groups. Eight panelists from each grade (24 in all) were assigned to the JIC group and the rest of the panelists (32) were assigned to the IDC procedure. Panel members in each group were approximately equivalent with respect to their demographic and professional characteristics.

Item Difficulty Categorization Procedure

Ten panelists participated in the IDC procedure for Grades 4 and 8, and 12 for Grade 12. This procedure addressed the question of whether specific statements about what students should know and be able to do in the achievement levels descriptions are supported by student performance on NAEP reading items. Each item in the item pool for a particular grade level was classified as an item that students "can do," students "can't do," or that students would find "challenging." The classifications of "can do," "can't do," and "challenging," were statistically determined according to the following rules.

"Can Do" Items

The "can do" category included items with at least a 50% probability of correct response at the *lower* bound of the achievement level. This ensured that the probability of correct response to those items for students who scored above the lower bound of a particular achievement level was greater than 50%, while students scoring below the lower bound had a probability of correct response of less than 50%. This definition provides a minimum test for what students within each level "can do."

"Can't Do" Items

The "can't do" category included items with less than a 50% probability of correct response at the *upper* bound of a particular achievement level category. These items had less than a 50% probability of a correct response for *all* students in that achievement level—even those at the upper bound. *Note: Because there is no upper bound for the Advanced level, only those items for which the probability of correct response at the lower bound of Advanced was less than 50% were excluded.*

"Challenging" Items

The remaining items ("challenging") were in neither of the previous two categories. These residual items have the greatest potential to exemplify what students at a particular achievement level "should do." The achievement levels descriptions cover a range of performances demarcated by the lower bound and upper bound scores. (The "can do" items were classified on the basis of lower borderline student performance.) The challenging items are more likely to correspond to the *level* descriptions than are the "can do" items because the latter are "challenging" for students performing at the next *lower* level. The probability of correct response to these items was greater than or equal to 50% at *some score point(s) within* the achievement level range. The probability of correct response for students scoring at higher achievement levels would be greater than 50%. Some students within the particular achievement level and all students at lower achievement level(s) had less than 50% probability of correct response. This category thus included *some* items that *some* students within the particular achievement level *could do* and *some* items that *some* students within this achievement level *could not* do. These items can be targeted to compare to the ALDs which do, in fact, cover a range of knowledge and skills that students *should know and be able to do.*

Panelists examined items in the "can do" and "challenging" categories at each achievement level to determine whether the knowledge and skills required for the items corresponded to the statements included in the ALDs. (Please see Appendix E for an example of the IDC lists.) Similarly, panelists examined the "can't do" items for each achievement level to determine whether descriptive statements indicated that students *should* know and be able to do things that were included in the "can't do" tasks. Panelists were asked to compile a list of the inconsistencies they observed to share with other panelists in the group. Panelists were asked to report specific aspects of the ALDs that seemed to lack support on the basis of the IDC procedure. Finally, they were asked to make

recommendations based on their conclusions. These recommendations were to be shared with the panelists participating in the JIC procedure when both groups had completed their tasks.

Judgmental Item Categorization Procedure

In this procedure, panelists were asked to classify each item according to the descriptions of achievement levels. Panelists were instructed to classify items into the achievement level that best described the knowledge or skill required to correctly respond. For example, an item was classified into the Basic level if panelists thought it matched the description of performance at the Basic achievement level. If the item seemed to match more than one ALD, the panelists were instructed to classify it at the lowest level of achievement. This "lowest level" rule reflects the cumulative nature of student achievement. That is, knowledge and skills associated with Basic level performance are assumed to hold for performance at both the Proficient and Advanced levels. Items that did not match any of the ALDs were put in a "Can't Classify" category. Although there is no description for the Below Basic level, it is an implied level and panelists were instructed to classify items as such. Panelists recorded their classification on a special form. (Please see Appendix F for the JIC form.) Panelists classified items in the order that they appeared in a block. Open-ended items that were scored polytomously (i.e., items that were scored for partial credit) with m levels were considered m-1 items.

The first classifications were made by each panelist working independently. After finishing a block of items, panelists were instructed to select a partner with whom they would "adjudicate" their classifications and try to reach agreement. Panelists were instructed to select different partners for adjudicating each block. When all items had been classified, panelists worked in their grade group to reach agreement on classifications. Panelists were instructed to note items for which agreement could not be reached. Consensus was a goal—not a requirement—for these panelists. Upon completion of the task, panelists were asked to evaluate the composite of items classified within each achievement level to determine the extent to which the ALDs were consistent with the skills and knowledge covered by the assessment. In particular, items in the "can't classify" category were to be analyzed to determine whether the items represented segments of the framework that had not been included in one or more achievement level descriptions. If so, this would signal a need to change one or more descriptions. If not, it would signal a misalignment of the item pool with respect to the framework. Similarly, if panelists found that only a few aspects of one or more ALDs were assessed, this should be reported as a signal of the lack of evidence to support reporting results in terms of achievement levels.

Joint Evaluation of the ALDs by Grade Groups

After the panelists assigned to each group completed their respective evaluations, they developed a group report to share with the other panelists in their grade level group. They then jointly evaluated the ALDs to determine whether the

descriptions were appropriate for reporting student performance on the NAEP. The panelists' goal was to reach a general conclusion on the recommendations regarding the use of the achievement levels for reporting NAEP results in 1994 and subsequent assessments using the current assessment framework. Panelists were asked to recommend whether any specific aspects of the ALDs should be deleted, added, or modified. They were also asked whether any misalignments observed appeared to be because the achievement levels' cutscores should be adjusted. Finally, as part of the grade level evaluation, panelists were asked to examine the ALDs for the other grades to determine whether the skills and knowledge levels were consistent with the assessment for a particular grade level.

Results

Three sets of results are presented. The group evaluations of the achievement levels are presented, the grade level results are presented, and results of an analysis of the outcomes of the judgmental classifications (JIC) relative to the statistical classifications (IDC) are presented. The results of the IDC procedures are presented as panelists' observations of the correspondence between the descriptors and the performance of students on assessment items. The results of the JIC procedures are based on the analyses of the classification that panelists made of the assessment items. Grade level results are presented in the form of recommendations reached by all panelists in each grade group, based on the evaluations of ALDs by each of the two groups.

Results for the IDC Procedure

The IDC procedure was designed to answer the question *"Can students do what the ALDs say they should be able to do?"* To answer this question, panelists examined items in the "can do" and "challenging" categories at each achievement level to determine whether the knowledge and skills required for the items corresponded to the statements included in the ALDs.

In all three grade levels, the panelists observed that students were generally able to do what the ALDs indicated they should be able to do. There were certain exceptions to this, but panelists described these as "anomalies rather than trends."[2] Panelists generally expressed a rationale to explain why students could not do particular items that called for skills and knowledge specified in an achievement level description. Those explanations usually had more to do with peculiarities of the items than the skills and knowledge required to answer the items correctly. Thus, there were some items on the "can't do" lists that required knowledge and skill levels consistent with what students performing at that achievement level *should* be able to do. Panelists did not, however, conclude that these instances signalled a need to change the achievement levels. Rather, they

[2] From 12th grade reporter's notes on the IDC procedure.

generally attributed these to item structure, context effects, scoring rubrics, and similar item idiosyncrasies

Panelists in Grades 4 and 12 observed that there were tasks that students performing at the Basic level could do, but these tasks were not in the description of what students should know and be able to do. For the most part, these tasks involved making inferences[3].

Panelists in this group agreed upon the recommendations regarding the ALDs to share and discuss with panelists participating in the JIC procedure. Their joint recommendations are presented as grade level results.

Results for the JIC Procedure

Panelists working with the JIC procedure evaluated the descriptions *via* a judgmental item mapping. In this procedure, panelists were asked to classify each item according to the descriptions of the achievement levels. The classification of an item was based on each panelist's evaluation as to the lowest level of performance (Basic, Proficient, or Advanced) required to answer it correctly. The first round of classification was performed by each panelist working independently. When all items in a block had been classified, panelists were asked to work in pairs to reach agreements on classifications. Then, all panelists in the group were asked to work together to reach agreement. Each panelist was asked to record his/her classification of each item. The following results are based on the classifications of items from this procedure.

Based on the classifications recorded by the panelists at each grade level, the modal classification of each item was determined. The number of items classified at each level, based on the modal classification, is presented in Table 4. Each dichotomous item was counted as one item, while each polytomous item was counted as either two or three items, depending on whether the item had three or four levels, respectively. For example, an item with three levels (inappropriate, partial, appropriate) was counted as two items. "Inappropriate" responses were not classified, but a "partial" response could be classified at one achievement level while an "appropriate" response was classified at another level. This method accounted for a total of 100 items for Grade 4, 145 items for Grade 8, and 155 items for Grade 12.

Grade 4 panelists and grade 12 panelists classified all the items in either the Basic, Proficient, or Advanced level. Grade 8 panelists, however, classified some items in the Below Basic category and some in the "cannot classify" category. Grade 8 panelists classified most of the items at the Basic level.

---

[3] Grade 4 panelists also mentioned tasks involving "author's device" and "evaluation."

Table 4

Number of Items Classified at Each Achievement Level

Level	Grade		
	4	8	12
Below Basic	0	3	0
Basic	37	112	80
Proficient	46	23	64
Advanced	17	5	11
"Can't Classify"	0	2	0
Total	100	145	155

Panelists identified no specific aspects of any achievement levels descriptions that were not represented in the reading assessment. Panelists did comment, however, that there were *not enough items* requiring the level of skills and knowledge required for advanced level performance. Some noted that the items were not "rich" enough to elicit the responses required for advanced level performance. These observations are confirmed by evidence found in Table 4. Grade 8 panelists classified only 5 of the 145 items (3%) at the Advanced level. Grade 4 panelists classified 17 of the 100 items as Advanced, and Grade 12 panelists classified 11 of 155 as Advanced. Grade 8 panelists noted that there were items that called for "[making] connections about related texts," but none of the grade 8 achievement levels descriptions indicated that this was something that students *should* be able to do. They recommended that this be added.

Table 5 shows the average p-values of items classified at each achievement level. If the difficulty of items were perfectly related to the achievement levels, then the most difficult items would all be classified at the Advanced level, the moderately difficult items at the Proficient level, and the least difficult items would all be classified at the Basic (or below) level(s). No such relationship is expected, however. The panelists were not given student performance data for this task. Theirs was a criterion referenced task for which only the content of each item and the scoring rubric were judged relative to the ALDs to determine the relative level of difficulty and corresponding level of achievement required to correctly respond. Nonetheless, one would logically expect the items classified at the lower levels of achievement to be the relatively easier items (as measured by student performance) and items classified at the higher levels to be relatively more difficult. Higher p-values (percentages of students correctly answering the items) should be observed for items classified at lower achievement levels, and the p-values should decrease with increases in the level of achievement.

The p-values used for polytomous items were the percentages of students scoring at or above a particular item score point. For example, the p-value at level 2 for a polytomous item with three levels (1, 2, and 3) is the percentage of students who scored 2 or higher, and the p-value for that item at level 3 is the percentage of students who scored 3.

The data in Table 5 indicate that the panelists were able to make sound judgments when mapping the items to the achievement levels. For Grade 12, for example, the average p-value for the 80 items classified at the Basic level is 69; for the 64 items classified at the Proficient level, the average p-value is 56; and, for the 11 items classified at the Advanced level, the average p-value is 21.

In Table 6, the average conditional p-value of each item classified at each level is averaged for the level. Average conditional p-values are computed by averaging

Table 5

Average P-Values of Items Classified at Each Level

Level	Grade		
	4	8	12
Below Basic		69 [54, 85][1]	
Basic	58 [15, 93]	59 [4, 91]	69 [12, 97]
Proficient	54 [3, 90]	31 [2, 77]	56 [6, 94]
Advanced	32 [0, 78]	2 [0, 5]	21 [1, 82]

[1] Bracketed numbers are the minimum and maximum p-values for items classified at teach level. For grade 8, "54" was the lowest probability of correct response for the 3 items classified as Below Basic and "85" was the highest.

Table 6

Average Conditional P-Values of Items Classified at Each Achievement Level at the Interval of Each Achievement Level

Grade 4

Modal Classification (no. of items)	Average Conditional P-Value Across the Interval for:			
	Below Basic	Basic	Proficient	Advanced
Basic (37)	39 [5, 88]	64 [17, 98]	79 [29, 99]	93 [56, 100]
Proficient (46)	33 [0, 78]	56 [1, 97]	73 [4, 100]	89 [30, 100]
Advanced (17)	15 [0, 58]	34 [0, 82]	49 [0, 94]	69 [6, 99]

Grade 8

Modal Classification (no. of items)	Average Conditional P-Value Across the Interval for:			
	Below Basic	Basic	Proficient	Advanced
Below Basic (3)	21 [16, 30]	53 [37, 77]	76 [59, 95]	91 [81, 100]
Basic (112)	33 [0, 83]	61 [2, 98]	81 [14, 100]	93 [46, 100]
Proficient (23)	12 [0, 44]	28 [1, 80]	47 [5, 95]	72 [18, 99]
Advanced (5)	0 [0, 1]	1 [0, 3]	5 [0, 12]	23 [6, 38]

Grade 12

Modal Classification (no. of items)	Average Conditional P-Value Across the Interval for:			
	Below Basic	Basic	Proficient	Advanced
Basic (80)	41 [1, 92]	69 [8, 100]	82 [23, 100]	93 [51, 100]
Proficient (64)	25 [1, 73]	52 [3, 97]	69 [8, 99]	87 [22, 100]
Advanced (11)	6 [0, 46]	18 [0, 93]	32 [1, 99]	57 [8, 100]

the probability of correct response at several score points within the range of scores from the lower cutpoint of an achievement level to the upper cutpoint. The entries in Table 6 are the averages over average conditional p-value of items. The rows are average conditional p-values for the items classified as Basic, Proficient, or Advanced. At grade 4, for example, 37 items were classified at the Basic level. The average conditional probability of correct response for these 37 items at points below the Basic level is 39%. The average conditional probability of correct response for those 37 items for students scoring from the lowest cutscore for Basic to the highest cutscore for Basic is 64%. And, the average conditional probability of correct response for those 37 items for students scoring from the lowest to the highest cutscore for Proficient is 79%.

The conditional p-values should increase across increasing levels of achievement because the probability of a correct response should increase as the assessment score increases. In general, the higher the student's performance score, the higher the probability of answering an item correctly. This general pattern is observed in Table 6, and it is another indication that panelists were able to make sound judgement when mapping items to achievement levels. Further confirmation of this is found when looking down column headings for each grade level. With the exception of the three items classified at the Below Basic level by 8th grade panelists, the average probability of correct response within an achievement level decreases as the achievement level for the item classification increases. The 3 items classified as Below Basic by 8th grade panelists had, on average, only a 53% probability for students scoring throughout the range of Basic achievement. Their high expectations are further evidenced by the 5 items classified as Advanced. The average probability of correct response for those items is only 23% for students scoring throughout the top ranges of the scale from the lowest cutscore of Advanced and higher. The highest average probability of correct response for the 5 items was 38%. Further investigation revealed that the 5 items are, in fact, all constructed response items and only the highest score (3 for items scored 1-3 and 4 for items scored 1-4) was identified as corresponding to Advanced level performance by these panelists.

This evidence shows that the achievement levels descriptions can be used to delineate the level of achievement required for successful performance on the many exercises and tasks on the NAEP. With the exception of the rather stringent item classifications by the 8th grade panelists, the average conditional p-values range from around 60% to 70% for level at which the items were classified.

Panelists who participated in the JIC procedure were asked to make recommendations regarding the achievement levels descriptions to be shared with other panelists in their grade group who had participated in the IDC procedure. These recommendations are summarized in the following section.

Joint Evaluation Results (Recommendations and Observations)

After panelists completed their respective procedures, they jointly evaluated the achievement levels. *For all three grade levels, panelists* ***supported*** *the achievement levels that were set in 1992.* They recommended, however, that the 1992 achievement levels be used in reporting NAEP reading results with some modifications—both editorial and substantive.

All grade level panels recommended that the ALDs be presented in a more "reader friendly" format (e.g., "bulleted" text or matrix format). Panelists representing Grades 4 and 8 provided examples of how, in their opinion, the ALDs should be presented. (Please see Appendix G.) The substantive changes are more grade specific and are presented below according to grade level[4]. These recommendations and observations are those presented to the whole group session at the conclusion of the Reading Revisit. Some recommendations and observations that they reported do not pertain to the ALDs, but they are presented here as well to provide as clear a picture as possible of the panelists' evaluations.

[4] The grade level recommendations are not in the same format. They are reproduced here to reflect, as closely as possible, the manner in which their respective recommendations were presented.

Grade 4

1. Basic descriptors do not reflect what the Basic level students "can do." (There were tasks that Basic level students could do that were not included in the descriptors.) For example, beginning with simpler forms and progressing to more sophisticated levels these students can: i.e., inference, author device, evaluation.

2. We would recommend that approaches skills, strategies generally are employed in reading literary text and informational text and should be indicated for both kinds of text/parallel construction.

3. Reevaluate the format of the descriptors. Make them more user friendly. For example, put in scope-and-sequence or chart form, or include one of those formats as an addendum.

4. Some advanced descriptors are not reflected in test items; especially the 4-level rubrics do not describe/require the highest levels included in "advanced" descriptors. (Some questions were not "rich" enough to illicit the response of Advanced level performance.)

The overall effort of this assessment, the descriptors, this process goes in the right direction and we support it.

1. Logical progression from basic through proficient to advanced generally present.

2. Inclusion of questions requiring higher level thinking.

Grade 8

The 8th grade panelists reviewed the descriptors and came up with the following conclusions:

Basic:

Clarify the term "theme."
(gist, main idea, recurring idea, or universal truth)
The word theme is misleading in the Basic Description because many people interpret theme as a literary device. Actually, students performing at the Basic level should be able to identify recurring ideas or major ideas. Use the phrase "main idea."

Line 8 [should read]: They should make simple inferences and draw conclusions supported by information in the text.

Students [performing] at the Basic level would not be expected to identify literary devices and elements.

Proficient:

Literary text — Eliminate the examples for literary devices and elements or include a more extensive list. (Example: Personification and foreshadowing add conflict, symbolism, and point of view.) The questions do not just reflect personification and foreshadowing.

Line 11 [should read]: "...with complex inferences based on the text."

Line 13 [should read]: "They should be able to interpret practical text and judge the importance..."

Add the following: "Students [performing at the Proficient level] should be able to make connections between related texts."

Advanced:

Add "make connections between related texts." There are items on the test that require students to do this. ("Cady" and "I Am One") and "Oregon Trail" and "Nettie").

Conclusions/Observations:

- ✓ We find correspondence between descriptors and performance on the questions.
- ✓ There should be parallel structure in [and] among all paragraphs (i.e., such as).
  Example: In Basic 244, eliminate the parenthesis and e.g. and use "such as" which is used in Proficient 283. In Advanced 328, eliminate the parenthesis and "i.e.," use "such as setting, plot, characters, and theme."

Grade 12

Recommendations:

1. The Achievement Levels [Descriptions] should be presented in a reader friendly format.

2. Include the compare and contrast skill at the Proficient level for inferences and Basic level for identification.

3. Include the word "explicit" at the Basic level, since "literal" is not a synonym of "explicit."

4. Remove "[such as] irony and symbolism" from the Proficient description.

5. At the Proficient level add the word "process [and]" before "apply" in working with practical text.

6. Include tasks/items on the use of personal experience to evaluate the usefulness of text information as indicated in the Proficient description for practical text.

7. Change the word "personal experience" to "background knowledge and experience" as seen in all levels.

8. In [the] Basic description remove the words "For example."

Observations:

1. Text needs to reflect multiple voices in literature.

2. Text length is excessive in many cases.

3. Time constraints and sustained reading requirements may create obstacles affecting achievement levels.

4. [There is a need to use] more appropriate texts for high school students throughout the nation.

5. Students are generally able to do what the achievement levels descriptions indicate they [should be able to] do, with certain exceptions which tend to be anomalies rather than trends.

6. Use a font which is compatible with readability and reader comfort.

7. Test tool does not draw on background experiences of all students (see point 4). (i.e., many students do not have experiences with bus schedules.)

Correspondence Between the IDC and JIC Procedures

Panelists assigned to work groups for the two procedures had a high level of agreement regarding the items. This finding held across the three grade levels. To determine the extent of IDC and JIC panelists' agreement, "hit rates" were calculated based on the results of the two procedures.

At a given achievement level, a "can do" item (classified as "can do" by the IDC procedure) is considered a "hit"[5] if it is classified by the JIC procedure at that level of achievement or lower. At that same achievement level, a "challenging" item is considered a "hit" if the item is classified by the JIC procedure at that level. Still at that same achievement level, a "can't do" item is considered a "hit" if it is classified by the JIC procedure at a higher level. Figure 1 illustrates this classification scheme. The shaded cells in the figure indicate the "hits" at each achievement level. For example, the Basic achievement level is shown at the top of Figure 1. If a "can do" item (first row of the table) was classified as either Basic or Below Basic by the JIC procedure, the classification is considered a hit (represented by the two shaded cells in the first row). This seems logical because the statistical classification (IDC) of the item is "can do" for the Basic level and the judgmental classification (JIC) of the item is "Basic." Thus, both classifications are the same. Similarly, items statistically classified in the IDC procedure as "challenging" at the Basic level and classified as Basic by the JIC procedure would represent a "hit" because challenging items are items that have at least a 50% probability of correct response within the range of Basic level achievement. Items statistically classified as "can't do" by the IDC procedure are those that even students at the highest score point in the Basic level have less than a 50% probability of correct response. If those items were classified by panelists in the JIC procedure as either Proficient or Advanced, this would also be considered a "hit."

Cross tabulations of items based on judgmental categorizations by the panelists and categorizations based on item difficulty are presented in Tables 7, 8, and 9. At each achievement level items were categorized, based on item difficulty, as "can do," "can't do," and "challenging." The table entries are the numbers of "can do," "can't do," or "challenging" items at each level that panelists categorized as Below Basic, Basic, Proficient, or Advanced.

Hit rates are shown in parentheses at the top of each table. For example, the hit rate in Table 7 for the Grade 4 Basic level is 48. Hit rates were calculated as follows: Within each achievement level, the numbers in the shaded cells (hits) were added together and then divided by the total number of items for that grade.

---

[5] Alternative definitions of "hit" are, of course, possible. The definition stated here is both logical and conservative. An additional scheme is presented in Appendix H.

Figure 1

Hits

Basic

Item Difficulty Category	Judgmental Category			
	Below Basic*	Basic	Proficient	Advanced
"Can Do"				
"Challenging"				
"Can't Do"				

Proficient

Item Difficulty Category	Judgmental Category			
	Below Basic*	Basic	Proficient	Advanced
"Can Do"				
"Challenging"				
"Can't Do"				

Advanced

Item Difficulty Category	Judgmental Category			
	Below Basic*	Basic	Proficient	Advanced
"Can Do"				
"Challenging"				

* Only for Grade 8 were any items classified as "Below Basic." The "Below Basic" category appears only on Table 8 for Grade 8.

Table 7

Cross Tabulation of Item Difficulty Categorization at Each Achievement Level (IDC) *vs.* Judgmental Item Categorization (JIC) for Grade 4*

Basic (48)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**21 (49/57)**	18 (42/39)	4 (9/24)	43
"Challenging"	**9 (28/24)**	19 (59/41)	4 (13/24)	32
"Can't Do"	7 (28/19)	**9 (36/20)**	**9 (36/53)**	25
Total	37	46	17	100

Proficient (79)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**30 (40/81)**	**37 (49/80)**	8 (11/47)	75
"Challenging"	4 (36/11)	**5 (45/11)**	2 (18/12)	11
"Can't Do"	3 (21/8)	4 (29/9)	**7 (50/41)**	14
Total	37	46	17	100

Advanced (93)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**34 (40/92)**	**42 (49/91)**	**10 (12/59)**	86
"Challenging"	3 (21/9)	4 (29/9)	**7 (50/41)**	14
Total	37	46	17	100

*The numbers in parentheses are percents (row/column).

**Hit Rate

Table 8

Cross Tabulation of Item Difficulty Categorization at Each Achievement Level (IDC) *vs.* Judgmental Item Categorization (JIC) for Grade 8*

Basic (80)**

Item Difficulty Category	Judgmental Category				Total
	Below Basic	Basic	Proficient	Advanced	
"Can Do"	**1 (2/33)**	**61 (94/54)**	3 (5/13)	0 (0/0)	65 (/45)
"Challenging"	1 (3/33)	**32 (86/29)**	4 (11/17)	0 (0/0)	37 (/26)
"Can't Do"	1 (2/33)	19 (46/17)	**16 (39/70)**	**5 (12/100)**	41 (/29)
Total	3 (2/)	112 (78/)	23 (16/)	5 (3/)	143

Proficient (81)**

Item Difficulty Category	Judgmental Category				Total
	Below Basic	Basic	Proficient	Advanced	
"Can Do"	**2 (2/67)**	**93 (91/83)**	**7 (9/30)**	0 (0/0)	102 (/71)
"Challenging"	1 (4/33)	15 (60/13)	**9 (36/39)**	0 (0/0)	25 (/17)
"Can't Do"	0 (0/0)	4 (25/4)	7 (44/30)	**5 (31/100)**	16 (/11)
Total	3 (2/)	112 (78/)	23 (16/)	5 (3/)	143

Advanced (92)**

Item Difficulty Category	Judgmental Category				Total
	Below Basic	Basic	Proficient	Advanced	
"Can Do"	**3 (2/100)**	**108 (85/96)**	**16 (13/70)**	**0 (0/0)**	127 (/89)
"Challenging"	0 (0/0)	4 (25/4)	7 (44/30)	**5 (31/100)**	16 (/11)
Total	3 (2/)	112 (78/)	23 (16/)	5 (3/)	143

*The numbers in parentheses are percents (row/column).

**Hit Rate

Table 9

Cross Tabulation of Item Difficulty Categorization at Each Achievement Level (IDC) *vs.* Judgmental Item Categorization (JIC) for Grade 12*

Basic (66)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**59 (69/74)**	26 (30/41)	1 (1/9)	86 (/55)
"Challenging"	**12 (43/15)**	16 (57/25)	0 (0/0)	28 (/18)
"Can't Do"	9 (22/11)	**22 (54/34)**	**10 (24/91)**	41 (/26)
Total	80 (52/)	64 (41/)	11 (7/)	155

Proficient (90)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**71 (62/89)**	**42 (37/66)**	1 (1/9)	114 (/74)
"Challenging"	7 (24/9)	**19 (66/30)**	3 (10/27)	29 (/19)
"Can't Do"	2 (17/2)	3 (25/5)	**7 (58/64)**	12 (/8)
Total	80 (52/)	64 (41/)	11 (7/)	155

Advanced (97)**

Item Difficulty Category	Judgmental Category			Total
	Basic	Proficient	Advanced	
"Can Do"	**78 (55/98)**	**61 (43/95)**	**4 (3/36)**	143 (/92)
"Challenging"	2 (17/2)	3 (25/5)	**7 (58/64)**	12 (8)
Total	80 (52/)	64 (41/)	11 (7/)	155

*The numbers in parentheses are percents (row/column).

**Hit Rate

The hit rate for the Grade 4 Basic achievement level was calculated by adding 21, 9, 9, and 9, and then dividing this quantity (48) by 100. The hit rate is the number of hits as a percentage of items.

Except for the Basic level at Grades 4 and 12, the hit rates were quite high, ranging from 79% (Grade 4 Proficient) to 97% (Grade 12 Advanced). For Grades 4 and 12 (Basic), the misses (i.e., non hits) were examined further[8] to determine whether they corresponded to items that required examinees to make inferences.

Recall that panelists in the IDC procedure for Grades 4 and 12 observed that items requiring inferences were classified as "can do" items for students performing at the Basic level, although this skill was not included in the Basic achievement level description. JIC panelists were instructed to use the achievement level descriptions to classify items. If these items required students to make inferences, then JIC panelists would most likely have classified the items at the Proficient level, since that is the lowest level for which "inferential" skills are specified for student achievement. The examination of the content of the 18 items in question for grade 4 Basic (the 18 items classified as Proficient by JIC panelists) revealed that 10 of the 18 items required the student to make some type of inference(s). The examination of items in question for Grade 12 revealed that 12 of the 26 items required students to make inferences. If the 10 items for Grade 4 had been classified as Basic instead of Proficient, the hit rate would have been 58%. If the 12 items at Grade 12 had been classified as Basic instead of Proficient the hit rate would have been 74%.

Overall, the hit rates found in Tables 7, 8, and 9 indicate a high correspondence between the IDC and JIC classifications. This suggests that panelists using their understanding of the assessment framework and the achievement levels descriptions *only* classified items in a way that was highly consistent with the statistical criteria used for classifying the items in the IDC procedure.

The convergence of the findings from the two procedures must be seen as clear evidence in support of the validity of the 1992 Reading achievement levels and in support of their use in reporting 1994, and subsequent, NAEP Reading results.

---

[8] David Duer, ACT reading specialist and content consultant for the 1992 Reading ALS process, evaluated the items for this purpose.

Conclusion

Two procedures were implemented to evaluate the 1992 NAEP Reading Achievement Levels in this "revisit." In the IDC procedure, panelists evaluated the achievement levels with respect to items that had been categorized according to statistical criteria. Panelists were to work from the item mappings provided to them and determine whether the achievement level descriptions of what students should know and be able to do corresponded to the item categorizations based on student performance on the NAEP, i.e. what students *can* (did) do. For this procedure, the item classifications were given to panelists and they were used to evaluate the descriptions.

In the JIC procedure, on the other hand, panelists used the achievement levels descriptions to develop the item mappings or categorizations. They then evaluated the achievement levels descriptions on the basis of the clarity, meaning, and utility of the descriptions for categorizing measurement indicators of the knowledge and skill levels required for each achievement level.

In the IDC procedure, the actual performance of students at each achievement level was used to categorize items; in the JIC procedure, the panelists' judgements of student performance were used. The high correspondence between the judgmental classification and the statistical classification provides compelling evidence that the achievement levels descriptions communicate clearly and accurately with respect to student performance. Further, the recommendations developed by the panelists involved in the two evaluation methods were quite similar. The fact that the recommendations made by the two sets of panelists were similar and confirming of the achievement levels seemed a sufficiently positive outcome to support the use of the achievement levels for reporting the 1994 results. The fact that the judgmental item mappings correspond so closely to the statistical item mappings greatly strengthens this conclusion.

To the extent that panelists recommended changes in the achievement levels descriptions, it was to increase the requirements for the Basic level descriptions, based on actual student performance. Contrary to this conclusion, however, was the earlier conclusion from the NAE evaluation that the levels were too high. The findings of this study cast doubt on the recommendations of the NAE panel that the achievement levels set in 1992 should not be used for reporting results on the Reading NAEP.

References

National Academy of Education (1993). *A report of the National Academy of Education Panel on the evaluation of the NAEP trail state assessment: An evaluation of the 1992 achievement levels.* Stanford, CA: National Academy of Education, Stanford University.

Appendix A

List of Observers, Panelists, and Staff

Reading Revisit
List of Staff and Observers

ACT Staff
Luz Bay, Facilitator
Don Gaudreau, Facilitator
Susan Loomis, Director
Charles Peters, Content Staff
Bishop White, Facilitator

NAGB Staff
Mary Lyn Bourque
Daniel Taylor

Observers
George Bohrnstedt, AIR
Jay Campbell, ETS
Peggy Carr, NCES
John Donoghue, ETS
Sharif Shakrani, NCES
Fran Stancavage, AIR

Reading Revisit Panelists

Teachers

Molly Chun
Applegate Elementary
Portland, OR

Grace Herr
West Linn High School
West Linn, OR

Carol Case
Mirabean B. Lamar Senior High
Houston, TX

Ruby Mayes
S.P. Waltrip Senior High
Houston, TX

Winfrey Bates
Mannsville Elementary School
Mannsville, KY

Roberta Horton
Custer County District H.S.
Miles City, MT

Pam Diamond
Hellgate Middle School
Missoula, MT

David Colburn
Flathead High School
Kalispell, MT

Donna Miller
Chinook High School
Chinook, MT

Katie Burnham
Pa Wau Lu Middle School
Gardnerville, NV

Cheryl Miller
Buchanan Elementary
Baton Rouge, LA

Terrence Smith
Verona School
Battle Creek, MI

Sarah Herz
Coleytown Middle School
Westport, CT

Pamela Perryman
Selah Middle School
Selah, WA

Joyce Boone
John Strange Elementary
Indianapolis, IN

Esther Dunnington
Grandview High School
Grandview, MO

Cara Terry
Lakewood High School
St. Petersburg, FL

Barbara Watson
Agricola Elementary School
Lucedale, MS

Lorraine Gerhart
Elmbrook Middle School
Elm Grove, WI

Theresa Lowe
Rancho Viejo School
Yuma, AZ

John Morrissey
Huntley Project Elementary
Worden, MT

Lynn Minderman
Honeoye Falls-Lima Central SD
Honeoye Falls, NY

Sandra Forsythe
Green Valley High School
Henderson, NV

Cynthia Freeman
Maryville High
Maryville, TN

Sarah Williams
Maryville Middle
Maryville, TN

David Fredette
Westborough High School
Westborough, MA

Ruth Johnson
Holmes High School
Covington, KY

Connie Clayton
Franklin High School
Franklin, WV

Caroline Downs
Worland Middle School
Worland, WY

Pamela McNair
Lemon G. Hine Jr. High School
Washington, DC

James Thompson
Simpson-Waverly School
Hartford, CT

Brenda Creel
Jessup Elementary
Cheyenne, WY

Beverly Hoffmaster
Berkeley Heights Elementary
Martinsburg, WV

Jody Alexander
Madison No. 1
Phoenix, AZ

Nonteacher Educators

Patsy Turner
Great River Co-operative
West Helena, AR

Kathleen Sanders
Los Angeles Unified S.D.
Wilmington, CA

Toni Walters
Oakland University
Rochester, MI

Susan Hodgin
Moscow School District
Moscow, ID

Faith Stevens
Haslett Public Schools
Haslett, MI

Daniel McQuagge
Delta State University
Cleveland, MS

Lory Johnson
Department of Education
Des Moines, IA

Maria Valeri-Gold
Georgia State University
Altanta, GA

Rita Gallagher
Unknown
Roswell, NM

Janet Williams
Bluewell Elementary School
Bluefield, WV

Roseine Church
Unknown
Cheyenne, WY

Sally Hellman
Unknown
Las Vegas, NV

Florence Wakuya
State Department of Education
Honolulu, HI

Bill Hammond
Department of Education
Atlanta, GA

Helen Schotanus
Department of Education
Concord, NH

Philip Yeaton
Unknown
Concord, NH

Martha Carter
Milwaukee Public Schools
Milwaukee, WI

Linda Brooks
Alcorn County School District
Corinth, MS

Evelyn Alford
East Baton Rouge School System
Baton Rouge, LA

Katie Young
State Department of Education
Baton Rouge, LA

Robert McKean
Havre Public Schools
Havre, MT

Richard Telfer
Univ. of Wisconsin-Whitewater
Whitewater, WI

Appendix B

Nomination Material

September 13, 1994

drmrms~ firstname~ lastname~
title?~
organization?~
address?~
citystatezip~

Dear drmrms~ lastname~:

We need your help in a very important matter. In 1992, The National Assessment Governing Board set achievement levels for the National Assessment of Educational Progress (NAEP) in Reading. This was the first time that the new NAEP Reading Framework had been used for the assessment, and the first time achievement levels had been set and used for reporting the Reading NAEP results. As is often the case for new programs, the National Assessment Governing Board (NAGB) wants to take a second look at—to "revisit"—the achievement levels before using them again to report results. The Reading NAEP was administered again this year, and the NAGB has asked us to review the achievement levels before the results of the 1994 Reading NAEP are reported. We need your help in identifying qualified people to serve on the review panels.

Our plan is to select panelists for this review who are either 4th, 8th, or 12th grade classroom teachers, or who are educators not teaching at the K-12 level. The panelists should be a broadly representative cross-section of the U.S. The teachers should comprise approximately sixty (60) percent of the panels and the nonteacher educators the remaining forty (40) percent.

We would greatly appreciate your submitting nominations of persons in your state who are qualified to serve as panelists. I have enclosed **sample** materials that were used for recruiting the panelists in the 1992 achievement levels-setting (ALS) process. Please note that the review process to which you will be nominating panelists is **not** the same as the achievement levels-setting (ALS) process. Those **sample** materials have been included to inform you about the qualifications required for the ALS panelists. I have also enclosed forms for **your** use in nominating the candidates. It is very important that these panelists have expertise in teaching and assessing reading. Please be sure to indicate the specific attributes or experiences that make the nominee "outstanding" for this task.

I realize that this is a ***very*** busy time for you, but we need your nominations as soon as possible! The meeting is scheduled for October 14-16, 1994, so we must have your nominations before the end of the month in order to draw the panels and make the arrangements for each panelist. The meetings will be at the Ritz-Carlton Hotel in St. Louis, Missouri. We will book airline reservations and send tickets to panelists.

The lodging cost will be billed directly to our master account, and two meals per day will be served as part of the meetings. In short, panelists will need their own money only for incidentals and one meal per day, plus parking and ground transportation. We will reimburse panelists for all normal costs such as the meal, ground transportation, parking, and so forth. In addition, we will reimburse the school/district for the cost of substitute teachers on the day or days the teachers must miss from class to participate in this study.

You may nominate as many panelists as you wish, although we have several criteria to be met that decrease the probability that *several* panelists from a single state would be selected to serve on the panel for any one grade level. We will select approximately 20 panelists for each of the three grade levels. We will attempt to draw our panels to include as many different regions of the U.S. as possible. Our first criteria, however, is that the panelists have expertise in reading and in the curricular requirements at the grade level to which they have been nominated to serve. Demographic diversity is also of greater importance than geographic diversity. So, if you know many teachers and nonteacher educators who would be outstanding panelists for this task, please nominate them.

We would be very happy to have you FAX the nominations to us **(319)339-3020** or mail them in the self-addressed, postage-paid envelope that was enclosed with the other material.

Thank you very much for your help. I have enclosed some information on NAGB and on the ALS process. If you have questions or concerns, please call me **(800)525-6929**.

Yours truly,

Susan Cooper Loomis, Ph.D.
Director, NAEP ALS Project
Research Division

Enclosures

GRADE 4
READING REVISIT
Teachers

a. Name _____

Home Address _____

City/State/Zip _____

Home Phone ( ) _____

School Name _____

School Address _____

City/State/Zip _____

School Phone ( ) _____

b. Total Years Teaching Experience

- ☐ 5-9
- ☐ 10-14
- ☐ 15 or more

c. Total Years Teaching Subject

- ☐ 2-4
- ☐ 5-9
- ☐ 10 or more

d. Race/Ethnicity

- ☐ Caucasian/White, non-Hispanic
- ☐ African-American/Black, non-Hispanic
- ☐ Asian-American or Pacific Islander
- ☐ American Indian/Alaskan Native
- ☐ Puerto Rican, Cuban, or Other Hispanic/Latino
- ☐ Other _____

e. Gender

- ☐ Male
- ☐ Female

f. Why do you feel this person is an outstanding candidate? (Please write on back or attach another page if more space is needed.)

_____

_____

_____

_____

_____

_____

_____

Please make as many copies of this form as you need.

RR

GRADE 4
READING REVISIT
Nonteacher Educators

a. Name _____

Home Address _____

City/State/Zip _____

Home Phone ( ) _____

Employer/Company Name _____

Work Address _____

City/State/Zip _____

Work Phone ( ) _____

b. Race/Ethnicity

- ☐ Caucasian/White, non-Hispanic
- ☐ African-American/Black, non-Hispanic
- ☐ Asian-American or Pacific Islander
- ☐ American Indian/Alaskan Native
- ☐ Puerto Rican, Cuban, or Other Hispanic/Latino
- ☐ Other _____

c. Gender

- ☐ Male
- ☐ Female

d. How is this person familiar with the subject matter and/or content area of Grade 4 reading? Why would this person be an "outstanding" panelist? Please use as much space as needed to provide this information.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Please make as many copies of this form as you need.

RR

Guidelines for Selecting Teacher Nominees and Information for Potential Nominees for Reading Panels

We ask that you nominate only those teachers whom you deem to be "outstanding." Outstanding teachers are those who are held in high regard by administrators, students, and/or fellow teachers, or who have been honored/recognized in some way, such as being named "teacher of the year." Teachers who have been very active in content-related professional associations, such as the National Reading Council (NRC), International Reading Association (IRA), or the National Council of Teachers of English (NCTE), would also be appropriate nominees.

We urge you to consult with your colleagues, with contacts you might have with local, state-level, or national content-related professional associations (e.g., NCTE or IRA), or with others who might assist you in identifying your best Reading teachers. We have also given your name to representatives of interested professional associations so that they might suggest names of teachers for you to consider.

Teacher nominees should have at least five (5) years of classroom experience, and must currently be classroom teachers. Two years of that experience (preferably most recent) should be at the 4th, 8th, or 12th grade levels, and should be in Reading (or disciplines with extensive and/or intensive reading requirements). Please nominate **up to** four (4) teachers for each grade level (4th, 8th, and 12th).

The Reading Achievement Levels-Setting process is scheduled for five days, August 21-25, 1992. The meetings are scheduled to begin on Friday evening in order to minimize the number of days people will need to be away from work. The meetings will be held at the Ritz-Carlton Hotel in St. Louis, Missouri, and all panelists will stay at the Ritz-Carlton from Friday night through Monday night. Panelists will not receive compensation for their participation *per se*, but their travel and lodging expenses will be paid and their meal costs will be reimbursed at the Federal *per diem* rate for St. Louis. In addition, we will reimburse the school (if appropriate) for the cost of hiring a substitute teacher for the days the teacher(s) selected are away from the classroom.

SAMPLE

We ask that you discuss this with the teachers that you might wish to nominate before you submit their names. We will assume that any teacher selected to serve as a panelist will have your permission to do so and to be away from their normal teaching responsibilities during the meeting period. It is important for you and any nominees to understand, however, that this is a nomination only. **Not all teachers nominated will be selected as panelists.** (In general, we expect to select only one or two teachers from any district.) We are requesting teacher nominations from three groups: superintendents, teacher association representatives, and private school principals. The final selection of teachers will be made in a way that will ensure that the panels are balanced with respect to gender, race/ethnicity, region of the nation, and other important characteristics. For that reason, we are asking you to identify the gender and race/ethnicity of teachers you nominate (see Nomination Form).

A pre-addressed, postage-paid envelope is enclosed for you to return your nominations. Please return your nominations at your earliest convenience, but please try to do so **by May 6, 1992**. In case you would prefer to FAX nominations, our FAX number is 319/339-3021.

Guidelines for Selecting Nonteacher Educator Nominees and Information for Potential Nominees on Reading Panels

Nominees must be educators (K-12, college/university, district/state level personnel) who are **not currently** classroom (K-12) teachers (other groups are being asked to nominate teachers). Nominees, for example, could include guidance counselors, curriculum specialists, principals, former teachers who are now administrators, college faculty members, college admissions officers, teachers of college freshmen, educational researchers, state commissioners of education, etc. Nominees should be knowledgeable about the learning and skills of students at the 4th, 8th, and 12th grade levels; knowledge of Reading skills would be particularly relevant.

We urge you to consult with your colleagues, with contacts you might have with local, state-level, or national professional associations (e.g., NCTE), or with others who might assist you in identifying outstanding non-teacher educators. We have also given your name to representatives of interested professional associations so that they might suggest names for you to consider.

Please nominate **up to** four (4) non-teacher educators for each grade level (4th, 8th, and 12th). **We encourage you to nominate yourself, if you so desire, and if you meet the specified criteria.**

The Reading Achievement Levels-Setting process is scheduled to last five days, August 21-25, 1992. The meetings are scheduled to begin on Friday evening in order to minimize the number of days panelists will need to be away from work. The meetings will be held at the Ritz-Carlton Hotel in St. Louis, Missouri, and all panelists will stay at the Ritz-Carlton from Friday night through Monday night. Panelists will not receive compensation for their participation *per se*, but their travel and lodging expenses will be paid and their meal costs will be reimbursed at the Federal *per diem* rate for St. Louis.

SAMPLE

We ask that you discuss this with the individuals you might wish to nominate before you submit their names. If necessary, we can contact the supervisors of nominees who are selected as panelists to secure permission for panelists to participate and to be away from their normal work responsibilities during the meeting period. It is important for you and any nominees to understand that this is a nomination only. **Not all nominees will be selected as panelists**. We are requesting nominations from other individuals throughout the nation. The final selection of panelists will be made in a way that will ensure that the panels are balanced with respect to gender, race/ethnicity, region of the nation, and other important characteristics. For that reason, we are asking you to identify the gender and race/ethnicity of nominees, including yourself (see Nomination Form).

A pre-addressed, postage-paid envelope is enclosed for you to return your nominations. Please return your nominations at your earliest convenience, but please try to do so **by May 6, 1992**. In case you would prefer to FAX nominations, our FAX number is 319/339-3021.

The National Assessment of Educational Progress (NAEP) and the Achievement Levels-Setting (ALS) Process

The NAEP

The National Assessment of Educational Progress (NAEP), an official U.S. Department of Education program, has provided information on the achievement and performance of students in the U.S. for over two decades. For each assessment, a nationally representative sample of approximately 35,000 to 100,000 students drawn from three age or grade levels has taken tests in various subject areas. The resulting data on student knowledge and performance have been accompanied by descriptive information allowing analysis of a variety of student experiences and background factors that correlate with student achievement.

The assessments have been designed to allow comparisons of student performance over time and among sets of students, grouped by region, type of community, race/ethnicity, and gender. The NAEP, commonly referred to as "The Nation's Report Card," is the most comprehensive and **only** continuing, valid source of information on what U.S. students know and can do, and how their performance has varied over time. A cornerstone of the Bush Administration's *Education 2000* agenda, the NAEP is vitally important as a measure of our students' academic achievement.

Achievement Levels-Setting

Public Law 100-297 (1988) contained the National Assessment of Educational Progress Improvement Act. The NAEP Improvement Act created the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP and provided that NAGB's responsibilities include:

- "Taking appropriate actions to improve the form and use of the National Assessment; and

SAMPLE

- Identifying appropriate achievement goals for each age and grade in each subject area to be tested under the National Assessment."

By defining levels of appropriate achievement on the National Assessment, NAGB seeks to increase greatly the significance and usefulness of NAEP results to educators, policymakers, and the American public.

To carry out these responsibilities as specified in the NAEP law regarding appropriate achievement goals, NAGB released a Request for Proposal on July 15, 1991 for setting achievement levels on the 1992 NAEP in Mathematics, Reading, and Writing. After reviewing several proposals, a contract was awarded to American College Testing (ACT) to design and administer a process that would allow NAGB to establish achievement levels on the NAEP to specify what students **should** know and be able to do. These levels will be determined in accordance with the policy framework, definitions, and technical procedures in NAGB policy titled, Setting Appropriate Achievement Levels for the NAEP, dated May 10, 1990.

Three achievement levels with clear distinctions between them will be established for each grade and subject tested under NAEP. These levels will be called:

Advanced: This higher level signifies superior performance beyond proficient grade-level mastery.

Proficient: This central level represents solid academic performance for each grade level tested.

Basic: This level, below proficient, denotes partial mastery of knowledge and skills that are fundamental for proficient work at each grade level tested.

It is NAGB's intention to use this framework of Basic, Proficient, and Advanced achievement levels as the primary means of reporting results for all newly developed assessments in 1992 and thereafter. The system is in contrast to NAEP's past practice of simply describing how students perform with no reference to standards of how well they ought to do.

The process of determining achievement levels will be a logical continuation of the national consensus effort used in developing the content and objectives of the NAEP. A broadly representative group of panelists will assist in defining the achievement levels using a proven judgment procedure to recommend levels of Basic, Proficient, and Advanced in terms of the NAEP subject areas.

As part of their deliberations, the panelists will prepare detailed descriptions of the subject-matter knowledge and skills proposed for each achievement level. These descriptions will be illustrated by representative sample items and scoring protocols. In preparing descriptions of achievement levels and assigning test items to them, panelists will use their best judgment and expertise and will also take into account a wide range of background information and frames of reference provided by ACT.

If you, or potential nominees, desire additional information about the NAEP or the ALS process, feel free to contact Dr. Mel Webb (ACT Project Director, 319/337-1472) or Dr. Susan Loomis (ACT Assistant Project Director, 319/337-1048).

A Brief Summary of Panelists' Responsibilities

Approximately two weeks prior to the meeting, all panelists will receive a packet that will include background and training materials, as well as a tentative agenda, for their review prior to the meeting. Panelists should be aware that every effort is made to assure their comfort and enjoyment during the ALS Meetings, but the schedule is rather rigorous and there is little free time once the meetings have begun. Panelists interested in visiting and sight-seeing in St. Louis should consider adding time before or after the scheduled meetings. We will attempt to accommodate those wishes when booking airline reservations for panelists. There will be 20-22 panelists for each grade-level (4th, 8th, and 12th). Approximately 55% will be current classroom teachers, 15% will be non-teacher educators, and 30% will be from the general public.

On the evening of Day 1, panelists will receive an overview of the Achievement Levels-Setting Meetings and the agenda, the National Assessment Governing Board (NAGB), and the NAEP Assessment program. On Day 2, they will be given an explanation of the content area framework with which they will be working, and they will begin training in other aspects related to the NAEP Assessment. They will then take generic definitions of Basic, Proficient, and Advanced (see the 2 page National Assessment of Educational Progress (NAEP) and Achievement Levels-Setting (ALS) Process description attached) and, in small groups, operationalize the definitions under ACT's and NAGB's guidance. Operational definitions will indicate the skills and abilities that students at a particular grade level (4th, 8th, or 12th) **should** have in order to be classified as performing at the Basic, Proficient, or Advanced level of achievement on the NAEP. Each panelist will also complete one form of their grade-level NAEP Assessment to familiarize them with the content and format of the test.

SAMPLE

On Day 3, panelists will discuss and modify their operational definitions of Basic, Proficient, and Advanced; reach agreement on common definitions for their grade-level; and receive copies of the operational definitions for later use. Panelists will then receive extensive training in their primary task for the achievement levels-setting process (see below), including practice carrying out that task. Day 3 will end with panelists performing Round 1 of the task.

The Task. Panelists' primary responsibility for setting achievement levels will be to examine individual items (test questions) for their grade-level NAEP Assessment. The panelists will be determining, for example, how difficult a 12th grade item would be for 12th graders. Then, with the operationalized definitions of Basic, Proficient, and Advanced in mind, panelists will decide what percentage of 12th grade students performing at the Basic level would get that item correct, what percentage of 12th grade students performing at the Proficient level would get that item correct, and what percentage of 12th grade students performing at the Advanced level would get that item correct. Panelists will repeat this procedure for a specified number of test items (the number varies by content area and grade-level). We anticipate that the **first** round of ratings will take 3½ to 4 hours. There will be 3 rounds of ratings, but subsequent rounds will not take as long as Round 1. At the end of the 3rd round of ratings, panelists' percentage correct estimates for each achievement level (Basic,

Proficient, and Advanced) for each item will be averaged to produce a percentage correct estimate for Basic, for Proficient, and for Advanced, over all the items. Each panelists' percentage correct estimate for Basic, Proficient, and Advanced will then be combined with the other 19-21 grade-level panelists' estimates, and averaged to produce a group estimate for Basic, for Proficient, and for Advanced. The procedure to be followed is referred to as the modified Angoff procedure.

On Day 4, panelists will first receive retraining in the rating procedure, information about the actual difficulty of items they rated, and information about the consistency of their ratings. They will then complete Round 2 of the item ratings, receive additional feedback about items and their ratings, then complete Round 3 of the rating process.

On Day 5, panelists will review their ratings and their operationalized definitions. We will discuss the entire process with panelists, and show panelists how their individual contributions helped produce a final result. We will then have panelists evaluate their experiences. We anticipate the meetings ending by early afternoon on Day 5.

Appendix C

Achievement Levels Descriptions & Procedures for Evaluating

Descriptions of Reading Achievement Levels for Basic, Proficient, and Advanced Student Performance on the 1992 NAEP

Preamble

Reading for meaning involves a dynamic, complex interaction between and among the reader, the text, and the context. Readers, for example, bring to the process their prior knowledge about the topic, their reasons for reading it, their individual reading skills and strategies, and their understanding of differences in text structures.

The texts used in the reading assessment are representative of common real world reading demands. Students at Grade 4 are asked to respond to literary and informational texts which differ in structure, organization, and features. Literary texts include short stories, poems, and plays that engage the reader in a variety of ways, not the least of which is reading for fun. Informational texts include selections from textbooks, magazines, encyclopedias, and other written sources whose purpose is to increase the reader's knowledge.

In addition to literary and informational texts, students at Grades 8 and 12 are asked to respond to practical texts (e.g., bus schedules or directions for building a model airplane) that describe how to perform a task.

The context of the reading situation includes the purposes for reading that the reader might use in building a meaning of the text. For example, in reading for literary experience, students may want to see how the author explores or uncovers experiences, or they may be looking for vicarious experience through the story's characters. On the other hand, the student's purpose in reading informational texts may be to learn about a topic (such as the Civil War or the oceans) or to

accomplish a task (such as getting somewhere, completing a form, or building something).

The assessment asks students at all three grades to build, extend, and examine text meaning from four stances or orientations:

- **Initial Understanding**—Students are asked to provide the overall or general meaning of the selection. This includes summaries, main points, or themes.
- **Developing Interpretation**—Students are asked to extend the ideas in the text by making inferences and connections. This includes making connections between cause and effect, analyzing the motives of characters, and drawing conclusions.
- **Personal Response**—Students are asked to make explicit connections between the ideas in the text and their own background knowledge and experiences. This includes comparing story characters with themselves or people they know, for example, or indicating whether they found a passage useful or interesting.
- **Critical Stance**—Students are asked to consider how the author crafted a text. This includes identifying stylistic devices such as mood and tone.

These stances are not considered hierarchical or completely independent of each other. Rather, they provide a frame for generating questions and considering student performance at all levels. All students at all levels should be able to respond to reading selections from all of these orientations. What varies with students' developmental and achievement levels is the amount of prompting or support needed for response, the complexity of the texts to which they can respond, and the sophistication of their answers.

Introduction

The following achievement-level descriptions focus on the interaction of the reader, the text, and the context. They provide some specific examples of reading behaviors that should be familiar to most readers of this document. The specific examples are not inclusive; their purpose is to help clarify and differentiate what readers performing at each achievement level should be able to do. While a number of other reading achievement indicators exist at every level, space and efficiency preclude an exhaustive listing. It should also be noted that the achievement levels are cumulative from Basic to Proficient to Advanced. One level builds on the previous levels such that knowledge at the Proficient level presumes mastery of the Basic level, and knowledge at the Advanced level presumes mastery at both the Basic and Proficient.

- *Description of Reading Achievement Levels for Basic, Proficient, and Advanced Fourth Graders*

Basic 212

Fourth-grade students performing at the basic level should demonstrate an understanding of the overall meaning of what they read. When reading texts appropriate for 4th graders, they should be able to make relatively obvious connections between the text and their own experiences. For example, when reading literary text, they should be able to tell what the story is generally about—providing details to support their understanding—and be able to connect aspects of the stories to their own experiences. When reading informational text, basic-level 4th graders should be able to tell what the selection is generally about or identify the purpose for reading it; provide details to support their understanding; and connect ideas from the text to their background knowledge and experiences.

Proficient 243

Fourth grade students performing at the proficient level should be able to demonstrate an overall understanding of the text, providing inferential as well as literal information. When reading text appropriate to 4th grade, they should be able to extend the ideas in the text by making inferences, drawing conclusions, and making connections to their own experiences. The connection between the text and what the student infers should be clear. For example, when reading literary text, proficient-level 4th graders should be able to summarize the story, draw conclusions about the characters or plot, and recognize relationships such as cause and effect. When reading informational text, proficient-level students should be able to summarize the information and identify the author's intent or purpose. They should be able to draw reasonable conclusions from the text, recognize relationships such as cause and effect or similarities and differences, and identify the meaning of the selection's key concepts.

Advanced 275

Fourth grade students performing at the advanced level should be able to generalize about topics in the reading selection and demonstrate an awareness of how authors compose and use literary devices. When reading text appropriate to 4th grade, they should be able to judge texts critically and, in general, give thorough answers that indicate careful thought. For example, when reading literary text, advanced-level students should be able to make generalizations about the point of the story and extend its meaning by integrating personal experiences and other readings with the ideas suggested by the text. They should be able to identify literary devices such as figurative language. When reading informational text, advanced-level 4th graders should be able to explain the author's intent by using supporting material from the text. They should be able to make critical judgments of the form and content of the text and explain their judgments clearly.

- *Description of Reading Achievement Levels for Basic, Proficient, and Advanced Eighth Graders*

Basic 244

Eighth grade students performing at the basic level should demonstrate a literal understanding of what they read and be able to make some interpretations. When reading text appropriate to 8th grade, they should be able to identify specific aspects of the text that reflect the overall meaning, recognize and relate interpretations and connections among ideas in the text to personal experience, and draw conclusions based on the text. For example, when reading literary text, basic-level 8th graders should be able to identify themes and make inferences and logical predictions about aspects such as plot and characters. When reading informative text, they should be able to identify the main idea and the author's purpose. They should make inferences and draw conclusions supported by information in the text. They should recognize the relationships among the facts, ideas, events, and concepts of the text (e.g., cause and effect and chronological order). When reading practical text, they should be able to identify the main purpose and make predictions about the relatively obvious outcomes of procedures in the text.

Proficient 283

Eighth grade students performing at the proficient level should be able to show an overall understanding of the text, including inferential as well as literal information. When reading text appropriate to 8th grade, they should extend the ideas in the text by making clear inferences from it, by drawing conclusions, and by making connections to their own experiences—including other reading experiences. Proficient 8th graders should be able to identify some of the devices authors use in composing text. For example, when reading literary text, students at the proficient level should be able to give details and examples to support themes that they identify. They should be able to use implied as well as explicit information in articulating themes; to interpret the actions, behaviors, and motives of characters; and to identify the use of literary devices such as personification and foreshadowing. When reading informative text, they should be able to summarize the text using explicit and implied information and support conclusions with inferences based on the text. When reading practical text, proficient-level students should be able to describe its purpose and support their views with examples and details. They should be able to judge the importance of certain steps and procedures.

Advanced 328

Eighth grade students performing at the advanced level should be able to describe the more abstract themes and ideas of the overall text. When reading text appropriate to 8th grade, they should be able to analyze both meaning and form and support their analyses explicitly with examples from the text; they should be able to extend text information by relating it to their experiences and to world events. At this level, student responses should be thorough, thoughtful, and extensive. For example, when reading literary text, advanced-level 8th graders should be able to make complex, abstract summaries and theme statements. They should be able to describe the interactions of various literary elements (i.e., setting, plot, characters, and theme); to explain how the use of literary devices affects both the meaning of the text and their response to the author's style. They should be able critically to analyze and evaluate the composition of the text. When reading informative text, they should be able to analyze the author's purpose and point of view. They should be able to use cultural and historical background information to develop perspectives on the text and be able to apply text information to broad issues and world situations. When reading practical text, advanced-level students should be able to synthesize information that will guide their performance, apply text information to new situations, and critique the usefulness of the form and content.

Description of Reading Achievement Levels for Basic, Proficient, and Advanced Twelfth Graders

Basic 269

Twelfth grade students performing at the basic level should be able to demonstrate an overall understanding and make some interpretations of the text. When reading text appropriate to 12th grade, they should be able to identify and relate aspects of the text to its overall meaning, recognize interpretations, make connections among and relate ideas in the text to their personal experiences, and draw conclusions. They should be able to identify elements of an author's style. For example, when reading literary text, 12th-grade students should be able to explain the theme, support their conclusions with information from the text, and make connections between aspects of the text and their own experiences. When reading informational text, basic-level 12th graders should be able to explain the main idea or purpose of a selection and use text information to support a conclusion or make a point. They should be able to make logical connections between the ideas in the text and their own background knowledge. When reading practical text, they should be able to explain its purpose and the significance of specific details or steps.

Proficient 304

Twelfth grade students performing at the proficient level should be able to show an overall understanding of the text which includes inferential as well as literal information. When reading text appropriate to 12th grade, they should be able to extend the ideas of the text by making inferences, drawing conclusions, and making connections to their own personal experiences and other readings. Connections between inferences and the text should be clear, even when implicit. These students should be able to analyze the author's use of literary devices. When reading literary text, proficient-level 12th graders should be able to integrate their personal experiences with ideas in the text to draw and support conclusions. They should be able to explain the author's use of literary devices such as irony or symbolism. When reading informative text, they should be able to apply text information appropriately to specific situations and integrate their background information with ideas in the text to draw and support conclusions. When reading practical texts, they should be able to apply information or directions appropriately. They should be able to use personal experiences to evaluate the usefulness of text information.

Advanced 348

Twelfth grade students performing at the advanced level should be able to describe more abstract themes and ideas in the overall text. When reading text appropriate to 12th grade, they should be able to analyze both the meaning and the form of the text and explicitly support their analyses with specific examples from the text. They should be able to extend the information from the text by relating it to their experiences and to the world. Their responses should be thorough, thoughtful, and extensive. For example, when reading literary text, advanced-level 12th graders should be able to produce complex, abstract summaries and theme statements. They should be able to use cultural, historical, and personal information to develop and explain text perspectives and conclusions. They should be able to evaluate the text, applying knowledge gained from other texts. When reading informational text, they should be able to analyze, synthesize, and evaluate points of view. They should be able to identify the relationship between the author's stance and elements of the text. They should be able to apply text information to new situations and to the process of forming new responses to problems or issues. When reading practical text, advanced-level 12th graders should be able to make a critical evaluation of the usefulness of the text and apply directions from the text to new situations.

October 7, 1994

DRMRMS~. FNAME~ LNAME~
ADDRESS~
CITY~, STATE~ ZIP~

Dear DRMRMS~. LNAME~:

Thank you for agreeing to serve as a panelist for the "Revisit" of the 1992 National Assessment of Educational Progress (NAEP) Reading Achievement Levels. This letter is to confirm your selection and to provide you with some written information about the study. The study will begin with participant registration at 4:30 P.M., followed by a social period at 5:00 P.M. and dinner at 6:00 P.M. The **Plaza Room** on the second floor of the hotel will be the location of these activities and the general orientation sessions following dinner.

Before giving you more information, however, I want to congratulate you for being nominated and selected to participate as a panelist. We asked nominators to nominate *outstanding* people to serve as panelists, and we selected the candidates who were most qualified and most outstanding. The credentials and experience of the panelists that you will be working with are impressive.

The study will be held in St. Louis, Missouri at the *Ritz-Carlton Hotel October 14-16*. We have made your room reservations at the hotel. Your airline tickets will be billed to our account and mailed to you. Your reservations **must** be made through Short's Travel; we **cannot** reimburse you for tickets purchased through another agent.

You will need money to cover your incidentals, including dinner Saturday night "on your own." We will provide dinner Friday night; and both continental breakfast and lunch will be provided on Saturday and Sunday. You will be reimbursed for normal allowable expenses, and those will include such costs as travel to and from your home to the airport, parking, and so forth. We are under contract to the National Assessment Governing Board, an agency of the U.S. Department of Education; and we are obliged to comply with government *per diem* expense restrictions. All necessary information for being reimbursed will be provided to you at the meeting. An expense report form is included in this material for you, and you will be reimbursed as soon as possible after the end of the meetings. Since we cannot pay you an honorarium for your services, we do make every effort to make your accommodations as pleasant and worry-free as possible and to minimize your "out-of-pocket" expenditures for this purpose.

Included in this mailing is a packet of advanced materials to help you prepare for the meetings. We attempt to provide all the information you need and to train you completely for the tasks that you will be performing at the meetings. We do not anticipate that this can be accomplished with the advanced materials we will send, but the advance material will help you get started in the process. Enclosed in this package is the *Summary Version of the Design Document* which describes the procedures for the 1994 achievement levels-setting (ALS) process. The process of selecting panelists and the process for setting achievement levels is similar to that used in the 1992 Reading ALS process. As you will see, the 1994 pilot study included several procedures and rating methods that were included for research purposes. Those were ***not*** part of the 1992 process, and only a couple will be implemented in the 1994 ALS process. This document ***does not*** exactly describe what was done for setting achievement levels in 1992 and it ***does not*** describe what you will be doing. What it does do is give you a good idea of what our ALS process is like. Most importantly, it will give you a good idea of the process by which the cutscores for the achievement levels were set.

An important difference between the 1992 ALS process and the 1994 process is the fact that preliminary achievement levels descriptions were ***not included*** in the Reading Framework document. Panelists had to develop the achievement levels descriptions before beginning the item rating process.

A copy of the achievement levels descriptions adopted by the National Assessment Governing Board (NAGB) and reported in the ***Nation's Reading Report Card*** are included in this packet of information, along with a copy of the ***Reading Framework***. The major focus of our work in the Revisit will be these achievement levels descriptions. In particular, we will address the correspondence between the numerical cutscores to represent student achievement on the NAEP scale and the descriptions of what students should know and be able to do.

Also included in this packet of materials are the address and telephone number of the Ritz-Carlton Hotel, along with information about where to find the hotel shuttle at the St. Louis airport. We have included information about St. Louis and some of the "free time" options that will be available for your enjoyment. Please note that the hotel requests that you not wear jeans and athletic attire in the lobby, restaurants, and "public" areas of the hotel. Ordinary office wear will be fine for these meetings.

I hope that you were told that no time is scheduled for religious services over the week-end. With so diverse a set of people serving as panelists, it is not possible to schedule time for religious observances of each. If this presents an unacceptable situation for you, please notify us that you must withdraw from the panel.

We know that some people have special dietary needs related to food allergies, diabetes, and so forth and that some people have physical conditions related to vision, sound, and mobility that require some special accommodation. The hotel will work with us to arrange to meet your needs to the greatest extent possible. It will be most helpful to have information about your special requirements well in advance of the meetings. If that information has not already been communicated, please provide us the needed information as soon as possible.

Finally, I want to be honest with you about the way you will spend your time. This is a very important task. You will have *little* free time. Our experiences with similar studies show that people made lasting friendships and valuable professional contacts. Because you will be working with other highly qualified persons with expertise and interest in reading, you will likely find the work to be at least somewhat pleasurable. We sincerely hope so!

Thank you again for agreeing to work with us as a panelist. We are looking forward to meeting you and working with you in St. Louis. Meanwhile, if you have questions or concerns, please call the Project Coordinator Teri Fisher, the Assistant Director Dr. Luz Bay, or me. Our toll-free number is 800/525-6929. Our "normal" office hours are 8:30 A.M. - 4:30 P.M. CDST, but we are often here later than that! Our FAX number is 319/339-3020.

See you soon!

Yours truly,

Susan Cooper Loomis, Ph.D.
Director, NAEP Project
Research Division

Procedures for Evaluating Achievement Levels
Set for the 1992 Reading NAEP

Panelists

A panel of approximately 20 members for each of the three grade levels assessed by NAEP (4th, 8th, and 12th) will be convened to revisit the achievement levels. Approximately 60% of the panel will be K-12 classroom teachers and 40% will represent other educators. Guidelines to nominators specified that panelists should have direct familiarity with knowledge and skills of students at the grade level, training and experience in teaching reading, involvement in activities of professional organizations, and so forth. The panelist were nominated by the coordinator or liaison of the state affiliates of the National Council of Teachers of English (NCTE) and the state reading specialists. A total of 165 nominators were identified from these two sources, and letters inviting nominations from them were mailed on September 13. Each nominator was invited to submit names of teachers and other educators for the panels.

Training

All panelists will be trained in the current NAEP Reading Framework, the NAGB policy definitions of the three achievement levels, and the 1992 Achievement Levels Descriptions. Further, all panelist will be trained in the background of NAEP, the development of the NAEP Reading Framework, roles of the principal NAEP actors (NCES and NAGB), the Achievement Levels Setting Process implemented in 1992, and the subsequent efforts commissioned by both NAGB and NCES. Finally, all panelists will be trained in the procedures to be implemented in this study.

Two procedures will be conducted to evaluate the achievement levels. About two-thirds of the panelists for each grade level will be involved in activities described for Group A and the remainder will participate in the activities described for Group B.

Group A Procedure

An evaluation of achievement levels descriptions, *via* a statistical item categorization, will be implemented using the **Item Difficulty Categorization (IDC)** procedure. This procedure addresses the question of whether specific statements in the achievement levels descriptions are supported by performance on items in the NAEP for students scoring within the ranges of the achievement levels. In this procedure, panelists will be given sets of items categorized for each achievement level. For each achievement level, each item will be categorized in one of three categories. Items in the "**can do**" category are those items with a probability of correct response of at least 50% at the **lower borderline** of the achievement level. Thus, the probability of correct response to those items for students performing **within** the achievement level is greater than 50%.

The "**can't do**" category is composed of those items that have less than a 50% probability of correct response at the **upper borderline** of the achievement level. These items have less than 50% probability for a correct response for **all** students **within** that achievement level.

Items that do not belong to either the "can do" or the "can't do" category are categorized as "challenging" items. These residual items have the greatest potential to serve as exemplary of what students at the achievement levels "should do." Note that because there is no upper borderline for the advanced level, items will be assigned to only two categories for this level—"can do" and "challenging".

Panelists will examine items in these categories at each achievement level to determine whether the knowledge and skills required for the items correspond to the statements included in the achievement levels descriptions. Similarly, panelists will examine the "can't do" items for each achievement level to determine whether descriptive statements are found to lack confirmation by student performance on some sets of items. Some statements will be found for which the item analyses will provide confirmation and some will perhaps be found for which there is a lack thereof. Panelists will be queried regarding their notion of the possible causes. Panelists will also be asked to report on the specific aspects of the descriptions that seem to lack support based on the item categorizations.

Group B Procedure

In procedure A, panelists are given sets of items that have been grouped into student performance categories. The second procedure aims to evaluate the descriptions *via* judgmental item mappings. In this procedure *panelists* will be asked to classify each item according to the descriptions of achievement levels. The classification of an item will be based on each panelist's evaluation as to the lowest level of performance (basic, proficient, or advanced) required to answer it correctly. The first round of classification will be made by each panelist working independently. When all items in a block have been classified, panelists will be asked to work in small groups to reach agreements on classifications. Then, all panelists in the group will be asked to work together to reach agreement. At each stage, panelists will note items for which agreement could not be reached. Note that consensus is a *goal* and *not* a requirement. Upon completion of the task, panelists will be asked to evaluate the items in each category with respect to the achievement levels descriptions and determine the extent to which the descriptions describe the skills and knowledge covered by the assessment.

Grade Level Procedure

When each group of panelists has completed their evaluations, they will be asked to develop a group report to share with the other half of the panelists in their grade group. They will then jointly evaluate the achievement levels descriptions to determine whether the descriptions are appropriate for reporting student performance on the NAEP. The ultimate goal is to reach a general conclusion on the achievement levels as a whole. Panelists will be asked to recommend whether any specific aspects of the achievement levels descriptions should be deleted, added, or modified. They will also be asked whether they would recommend adjusting the cutpoint for any of the achievement levels. Finally, as part of the grade level evaluation, panelists will be asked to examine the achievement levels descriptions for the other grades to determine whether the skills and knowledge levels and domains called for are aligned with the assessment for the grade level.

Recommendations

After all the evaluation procedures have been implemented, panelists will be asked for specific recommendations regarding the use of achievement levels descriptions and cutpoints for reporting the 1994 Reading NAEP results. No changes in the achievement levels descriptions nor cutpoints will be made by the panelist in this study. However, if panelists recommend changes, those changes will be fully documented and submitted to NAGB for further consideration.

Appendix D

Agenda

Agenda
Reading Revisit, October 14-16, 1994
Ritz-Carlton Hotel, St. Louis, Missouri

Friday, October 14

4:30 P.M. Check-In/Registration with ACT Staff in Promenade Room

Whole Group Session

5:00 P.M. Welcome and General Orientation Session: Plaza Room
Susan Loomis, ACT
Mary Lyn Bourque, NAGB

6:00 P.M. Social Period: Pavilion Room

6:30 P.M. Dinner: Pavilion Room

Grade Group Session

7:45 P.M.
Grade 4: Don Gaudreau, Plaza Room (tonight only)
Grade 8 Bishop White, Consulate Room
Grade 12: Luz Bay, Colonnade Room

Table Group Assignments
Take and Self-Score NAEP Exam

9:30 P.M. Adjourn

Saturday, October 15

Whole Group Session: Plaza Room

8:00 A.M. Continental Breakfast: Plaza Room Foyer

8:30 A.M. Review Process for Setting 1992 NAEP Reading Achievement Levels, Susan Loomis

9:15 A.M. Reach Understanding of NAEP Reading Framework and Achievement Levels (Basic, Proficient, and Advanced) for Grades 4, 8, and 12, Charles Peters

10:45 A.M. Break (Plaza Foyer)

(over)

Grade Group Session

11:00 A.M. (Grade 4: **Ambassador Room**; Grade 8 Consulate Room; Grade 12 Colonnade Room)
Train in Achievement Levels Descriptions: Internalize Achievement Level Descriptions

Whole Group Session

Noon Lunch: Pavilion Room

Grade Group Session

1:00 P.M. Continue Training in Achievement Levels Descriptions and Reach Understanding of Borderline Performance: Work with Items and Student Papers

2:15 P.M. Break

Group A **Grade Groups**: Breakout Rooms
2:30 P.M. Read Passages, Scoring Guides and Keys

Group B **Plaza Room**
2:30 P.M. Training in Judgmental Item Categorization Procedure; Begin Procedure

Group A **Colonnade Room**
5:30 P.M. Instructions on the Item Difficulty Categorization Procedure

6:30 P.M. Adjourn

Sunday, October 16

8:00 A.M. Continental Breakfast: Plaza Foyer

Group Sessions

8:30 A.M. **Group A Panelists: Grade Groups**
Begin Item Difficulty Categorization Procedure
(**Breaks as Needed**)

8:30 A.M. **Group B Panelists: Plaza Room**
Continue Judgmental Item Categorization Procedure
(**Breaks as Needed**)

NAEP Reading Revisit Agenda
Sunday, October 16
(continued)

Whole Group Session
Noon — Lunch: **The Grill** (Lobby Level)

Group Sessions
1:00 P.M.
- **Group A**: Arrive at Grade-Group Conclusions and Recommendations
- **Group B**: Arrive at Grade-Group Conclusions and Recommendations

Grade Group Sessions
2:00 P.M.
- Arrive at Grade Level Conclusions and Recommendations based on Group A and Group B Evaluations
- Provide Written Recommendations
- Group Discussion on Recommendations

Whole Group Session **Plaza Room**
4:00 P.M. — Overall Review of Recommendations; Wrap-up of Reading Revisit

4:30 P.M. — Adjourn

Appendix E

Example of Item Difficulty Characteristic List

Reading Revisit Study
Item Difficulty Categorization information for Grade 04

For Basic level:

"Can do" items:

001	3		SHORT DESCRIPTION OF ITEM CONTENT
001	7		SHORT DESCRIPTION OF ITEM CONTENT
001	8		SHORT DESCRIPTION OF ITEM CONTENT
001	9		SHORT DESCRIPTION OF ITEM CONTENT
001	10		SHORT DESCRIPTION OF ITEM CONTENT
001	11		
002	1		
002	2		
002	3		
002	4		
002	5		
002	8		
002	9		
002	10		
002	12		
003	2		
003	7		
004	8		
004	11		
005	1		
005	2		
005	7		
005	8		
005	10		
006	1		
006	2		
006	3		
006	7		
007	2		
007	4		
007	5		
007	6		
007	7		
007	8		
007	9		
007	10		
007	11		
008	1		
008	2		
008	9		
005	4	>=2:	
006	5	>=2:	
007	12	>=2:	

"Can't do" items:

001	1		SHORT DESCRIPTION OF ITEM CONTENT
003	3		SHORT DESCRIPTION OF ITEM CONTENT
003	5		SHORT DESCRIPTION OF ITEM CONTENT
003	9		SHORT DESCRIPTION OF ITEM CONTENT
004	1		SHORT DESCRIPTION OF ITEM CONTENT
004	2		
004	4		
006	4		
006	6		
007	1		
008	5		
008	6		
001	6	>=3:	
001	6	>=4:	

BEST COPY AVAILABLE

002	11	>=3:
003	1	>=3:
003	1	>=4:
004	7	>=3:
004	7	>=4:
005	4	>=4:
006	5	>=3:
006	5	>=4:
007	12	>=4:
008	8	>=3:
008	8	>=4:

Challenging items:

001	2	SHORT DESCRIPTION OF ITEM CONTENT
001	4	SHORT DESCRIPTION OF ITEM CONTENT
001	5	SHORT DESCRIPTION OF ITEM CONTENT
002	6	SHORT DESCRIPTION OF ITEM CONTENT
002	7	SHORT DESCRIPTION OF ITEM CONTENT
003	4	
003	6	
003	8	
004	3	
004	5	
004	6	
004	9	
004	10	
005	3	
005	5	
005	6	
005	9	
006	8	
006	9	
006	10	
007	3	
008	3	
008	4	
008	7	
008	10	
001	6	>=2:
002	11	>=2:
003	1	>=2:
004	7	>=2:
005	4	>=3:
007	12	>=3:
008	8	>=2:

For Proficient level:
"Can do" items:

001	2	SHORT DESCRIPTION OF ITEM CONTENT
001	3	SHORT DESCRIPTION OF ITEM CONTENT
001	4	SHORT DESCRIPTION OF ITEM CONTENT
001	5	SHORT DESCRIPTION OF ITEM CONTENT
001	7	SHORT DESCRIPTION OF ITEM CONTENT
001	8	
001	9	
001	10	
001	11	
002	1	
002	2	
002	3	
002	4	
002	5	
002	6	
002	7	
002	8	
002	9	
002	10	
002	12	
003	2	
003	4	
003	6	
003	7	
003	8	
004	3	
004	5	
004	6	
004	8	
004	9	
004	10	
004	11	
005	1	
005	2	
005	3	
005	5	
005	6	
005	7	
005	8	
005	9	
005	10	
006	1	
006	2	
006	3	
006	7	
006	8	
006	9	
006	10	
007	2	
007	3	
007	4	
007	5	
007	6	
007	7	
007	8	
007	9	
007	10	
007	11	
008	1	
008	2	
008	3	
008	4	

```
008    7
008    9
008   10
001    6       >=2:
002   11       >=2:
003    1       >=2:
004    7       >=2:
005    4       >=2:
005    4       >=3:
006    5       >=2:
007   12       >=2:
007   12       >=3:
008    8       >=2:

    "Can't do" items:
006    4              SHORT DESCRIPTION OF ITEM CONTENT
007    1              SHORT DESCRIPTION OF ITEM CONTENT
008    5              SHORT DESCRIPTION OF ITEM CONTENT
001    6       >=4:   SHORT DESCRIPTION OF ITEM CONTENT
002   11       >=3:   SHORT DESCRIPTION OF ITEM CONTENT
003    1       >=3:
003    1       >=4:
004    7       >=3:
004    7       >=4:
005    4       >=4:
006    5       >=4:
007   12       >=4:
008    8       >=3:
008    8       >=4:

    Challenging items:
001    1              SHORT DESCRIPTION OF ITEM CONTENT
003    3              SHORT DESCRIPTION OF ITEM CONTENT
003    5              SHORT DESCRIPTION OF ITEM CONTENT
003    9              SHORT DESCRIPTION OF ITEM CONTENT
004    1              SHORT DESCRIPTION OF ITEM CONTENT
004    2
004    4
006    6
008    6
001    6       >=3:
006    5       >=3:
```

For Advanced level:
"Can do" items:

001	1	SHORT DESCRIPTION OF ITEM CONTENT
001	2	SHORT DESCRIPTION OF ITEM CONTENT
001	3	SHORT DESCRIPTION OF ITEM CONTENT
001	4	SHORT DESCRIPTION OF ITEM CONTENT
001	5	SHORT DESCRIPTION OF ITEM CONTENT
001	7	
001	8	
001	9	
001	10	
001	11	
002	1	
002	2	
002	3	
002	4	
002	5	
002	6	
002	7	
002	8	
002	9	
002	10	
002	12	
003	2	
003	3	
003	4	
003	5	
003	6	
003	7	
003	8	
003	9	
004	1	
004	2	
004	3	
004	4	
004	5	
004	6	
004	8	
004	9	
004	10	
004	11	
005	1	
005	2	
005	3	
005	5	
005	6	
005	7	
005	8	
005	9	
005	10	
006	1	
006	2	
006	3	
006	6	
006	7	
006	8	
006	9	
006	10	
007	2	
007	3	
007	4	
007	5	
007	6	
007	7	

```
007     8
007     9
007    10
007    11
008     1
008     2
008     3
008     4
008     6
008     7
008     9
008    10
001     6       >=2:
001     6       >=3:
002    11       >=2:
003     1       >=2:
004     7       >=2:
005     4       >=2:
005     4       >=3:
006     5       >=2:
006     5       >=3:
007    12       >=2:
007    12       >=3:
008     8       >=2:

    Challenging items:
006     4               SHORT DESCRIPTION OF ITEM CONTENT
007     1               SHORT DESCRIPTION OF ITEM CONTENT
008     5               SHORT DESCRIPTION OF ITEM CONTENT
001     6       >=4     SHORT DESCRIPTION OF ITEM CONTENT
002    11       >=3     SHORT DESCRIPTION OF ITEM CONTENT
003     1       >=3
003     1       >=4
004     7       >=3
004     7       >=4
005     4       >=4
006     5       >=4
007    12       >=4
008     8       >=3
008     8       >=4
```

Appendix F

Example of Judgmental Item Categorization

Identification #: _____

Grade 8

Group B

Dorothea Dix

		Below Basic	Basic	Proficient	Advanced	Can't Classify
1						
2						
3	2=					
	3=					
	4=					
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						

Appendix G

Grade-Level Organizational & Format Suggestions for ALDs

Grade 4

	Initial Understanding (unreflected)	Developing Interpretation (more complete)	Personal Reflection and Response (text/persnl bkgrd)	Critical Stance (stand apart & consider text)
BASIC	• understand overall meaning • tell what story is gen about (lit) • tell what selection is gen about (info) • identify purpose for rdg info text	• provide details to support understanding (lit stories & info text)	• relatively obvious connections • connect aspects of stories to own experiences • connect ideas fr info text to own exp & knowledge	
PROFICIENT	• inferential & literal info • summarize story • summarize info • identify author purpose • key concepts	• inferential & literal info • making inferences, drawing conclus w/clear connections to text • conclus about char or plot • recognize relation (cause/effect, sim/diff)	• extend text ideas, make connections to own exp	
ADVANCED	• generalize about topics in selection, ie, point of story	• thorough answers, careful thought • ident. literary devices (ie, fig lang) • explain author's intent w/support fr text (info)	• integrate text, pers exp & other readings	• awareness of how authors compose & use lit devices • judge texts critically • critical judgments about form & content (make, explain clearly)

Grade 8
Descriptor Model

Proficient 283

Eighth grade students performing at the proficient level should be able to show an overall understanding of the text, including inferential and literal information. When reading text appropriate to eighth grade, students should be able to perform the following:

(1) Make clear inferences.

(2) Draw conclusions.

(3) Make connections to their own experience (including other reading experiences).

(4) Identify literary devices within the text.

Literary Text: Students should be able to:

- Give details or examples of themes identified.
- Use implied and explicit information in articulating themes.
- Interpret the characters' actions, behaviors, and motives.
- Identify the use of literary devices.

Informative Text: Students should be able to:

- Summarize the text using explicit and implied information.
- Support conclusions with inferences based on text.

Practical Text: Students should be able to:

- Describe text purpose(s).
- Support views with examples and details.
- Judge importance of certain steps and procedures.
- Make interpretations from text.

Grade 8

Situtations	Basic	Proficient	Advanced
Literary	• interprets literally • identifies themes • makes inferences • predicts logically	• makes connections to experiences • recognizes literary devices • makes inferences from explicit info • interprets characters' motives	• makes complex, abstract summaries & theme statments • describes the interactions of various literary elements • analyzes/evaluates the writing • explains how the use of literary devices affect both the meaning and response to the author's style
Informative	• identifies main idea • identifies author's purpose • makes inferences/ conclusions based on text • recognizes relationships (fact, ideas, events, concepts, cause/effect, chrono order)	• summarizes text drawn from text • supports conclusions w/inferences based on text	• extends text info by relating it to their experiences • analyzes the author's purpose and point of view • uses cultural and hist background info to develop perspectives on the text • applies text info to broad issues/world situations
Practical	• identifies main purpose • makes obvious predictions	• describes text purpose and supports view w/examples & details • evaluates steps/ procedures—and their practicality	• synthesizes info that will guide their performance • applies text info to new situations • critiques the usefulness of the form and content

Grade 12

Basic	Proficient	Advanced
Understanding some interpretation *relate* aspects to overall meaning *recognize* interpretations	Understanding *inferential* & literary	Describe more *abstract* ideas/ themes *analyze* meaning form *support* with texual examples *extend* info to exp & world
Make Connections relate ideas in text to own exper *draw* conclusions define elements of style explain theme	Make Connections *inferences to extend ideas* *adds connection from rdg* implicit inference should be clr be able to analyze author's lit dev	Responses thorough thoughtful extensive
Support Conclusions	Draw & Support Conclusions be able *to apply* info to spec sit interpret background info	Produce complex abstract summaries & theme statements drawn on cultural, hist, & personal info
Informative Text main idea use text info to support conclusion	Practical Text be able to apply info be able to use personal info to evaluate usefulness of text info	Be able to evaluate text info text *analyze* *synthesize* evaluate point of view identify relationship between author's stance & text be able to apply & evaluate value of info text
Make Logical Connections between text & own exper be able to explain purpose & significance of details or steps	Draw Inferences apply knowledge	

Appendix H

Alternative Method of Computing "Hits"

Alternative Method of Computing "Hits"

There were some questions about what constituted a hit with respect to the "challenging" items. Some believed that all item classified at a higher level by the JIC panelists should be counted as hits since they would definitely be challenging for students at the level under consideration.

We believe that the method presented in Figure 1 directly corresponds to the item classifications of the IDC methodology. Thus, that method seems the most direct test. Nonetheless, this method is presented for readers who are interested in alternatives.

The hit rate for the Advanced level is computed for "Can Do" items only because there are no "Can't Do" items for the Advanced level.

Grade 4:
Basic Hit Rate:
39 of 100 items in all = 39%
39 of 68 "Can Do" + "Can't Do" items = 58%

Proficient Hit Rate:
37 of 100 items in all = 37%
37 of 89 "Can Do" + "Can't Do" items = 83%

Advanced Hit Rate:
86% of all or 100% of "Can Do"

Grade 8:
Basic Hit Rate:
83 of 143 items in all = 58%
106 of 106 "Can Do" + "Can't Do" items = 78%

Proficient Hit Rate:
107 items of 143 items in all = 75%
107 of 118 "Can Do" + "Can't Do" items = 91%

Advanced Hit Rate:
89% of all or 100% of "Can Do"

Grade 12:
Basic Hit Rate:
91 items of 155 items in all = 59%
91 items of 127 "Can Do"+ "Can't Do" items = 72%

Proficient Hit Rate:
120 items of 155 items in all = 77%
120 items of 126 "Can Do" + "Can't Do" items = 95%

Advanced Hit Rate:
92% of all or 100% of "Can Do"